

**Republique algeria democratique and populaire.**  
**Ministry of Higher Education and Scientific Research.**  
**University 8 May 45 –Guelma-**  
**Faculty of Mathematics, Computer Science and Material Sciences.**  
**Computer Science department.**



**Master's Thesis**

**Sector:** Computer science

**Option:** Information and Communication Science and technology.

**Theme :**

---

**FarmlQ: AI-powered virtual assistant for precision livestock farming**

---

**Presented By :** *Messahel Ayoub , Kolli Marwa*

**Jury members :**

- **President :** Dr. Soussi Hakim
- **Supervisor :** Pr. Farou Brahim
- **Examiner :** Dr. Djakhdjakha Lynda
- **Professional pole :** Dr. Mermet Adila

**June 2025**

# *Thanks*

*Above all we thank God the Almighty who gave us the strength and courage to accomplish this modest work.*

*A big thank you to **Professor Farou Brahim** for his supervision, he was always available throughout the production of this dissertation, for the inspiration, help and time he was kind enough to devote to me.*

*To my families and friends for their encouragement, we were able to overcome all obstacles.*

*My sincere thanks also go to the jury members for the interest they showed in our project by agreeing to examine and judge our work.*

*My thanks also go to all the teachers at the university and also all the staff of the computer science department for their kindness.*

# *Dedications*

## *kolli marwa*

***To my beloved parents:*** For your unconditional love, unwavering support, and constant encouragement. Thank you for always believing in me, even when I doubted myself. This thesis is the result of your sacrifices, patience, and reassuring presence at every step. I dedicate it to you with all my heart.

***To the best person:*** Messahel Ayoub, my best friend, my teammate in this journey, and so much more. Since we were five years old, we've shared everything — joy, challenges, dreams, and now, the completion of this project. This work is more than just a thesis; it is the result of our bond, our persistence, and your unwavering courage. Thank you for always being there, for your constant support, and for the quiet strength that often carried me. I don't see this achievement as mine alone, but truly ours.

***my dear brothers :*** To my big brother Mouad, Thank you for your courage, your constant support, and your wise words. To my brother Mehdi, Even from a distance, you've never stopped being close to my heart. Your determination and courage in pursuing your dreams abroad inspire me every day. I'm proud of the path you're carving for yourself, and I know greater things await you.

***To my friends:*** Oumeddoure imene, Seddiki loubna, Bouzidi lina maya, medjaledi faiza, Malak, for their support and encouragement and their moral support and help.

***To myself:*** for believing, persevering, and never giving up. This work is a reflection of my commitment, my silent efforts, the sleepless nights, and the dreams I pursued with passion. I dedicate this project to myself in recognition of my determination, my resilience in the face of challenges, and the strength I found within when things got tough. I am proud of how far I've come. This thesis is not an end, but the beginning of a future I'm building with confidence.

And above all, to my mother None of this would have been possible without you. Your endless sacrifices, your patience, and your unconditional love are the true foundation of everything I accomplish. Thank you, from the bottom of my heart.

# *Dedications*

**Messahel Ayoub**

**To my parent:** whose unwavering support and unconditional love have been the pillars of my academic success. Your constant encouragement and sacrifices have been my source of inspiration throughout this journey. I dedicate this dissertation to you as a token of my eternal gratitude.

**ma précieuse binôme Marwa:** Thank you for being there throughout these five years, through the good times and the most difficult. Your support, your good humor, your perseverance, and your friendship have made all the difference. I am proud of what we have accomplished together.

**To my beloved sisters:** Kholoud and Soundes, for your unconditional love, quiet strength, and constant support— you have always been my anchor and my motivation.

**To my brother Aymen:** Thank you for always believing in me, even when I doubted myself. Your constant support, your quiet presence, and your encouraging words gave me the strength to keep moving forward.

**To my precious niece, Sadan:** a little star whose light fills my heart with joy and inspiration.

**To my friends and family:** who have encouraged and supported me throughout this adventure. Your presence, encouragement, and belief in my abilities have been inexhaustible sources of motivation. I dedicate this dissertation to you as a token of appreciation and gratitude for your constant support.

**And to my dear mother** you who have always been my refuge, my silent strength, and my greatest support. Your love, your prayers, and your caring presence have carried me throughout this journey. None of this would have been possible without you. I dedicate this success to you from the bottom of my heart.



---

# *Abstract*

Precision Livestock Farming is a system that introduces a welfare dimension while enhancing herd management through technological advances. Manual monitoring strategies which are a common practice in husbandry have time as well as cost constraints. This gap can permit health challenges to develop into key outbreaks within herds, resulting disruption of economic activities and the threat to public health, specially in the case of zoonoses. The innovative Precision Livestock Farming on the market developed within this framework, utilising A technologies, smart cameras and IoT sensors, offers solution for constant monitoring of the animals' health status, provide an early warning in the case of illnesses and allow giving appropriate support to every single animal. Economically, this assistant could offset some losses due to reduction in the ability to transmit diseases leading o reduced expenditure on not only veterinary treatment but also the number of deaths. The assistant's objectives focus on enabling self interactive and responsive system to health problems and behaviors in order to limit the number of human intervention and enhance the rate of response from a health perspective. The selected tools encompass individual tracking of animals with the use of smart cameras, the use of IoT sensors for the purpose of gathering environmental and behavioral context and using of deep learning algorithms for the purpose of predictive analysis. Key actions would be supported by data management platform for this purpose, at the same time notification systems would serve the purpose of rapid responses..

**Keywords :** AI, IoT, Smart Farming, monitoring, Deep Learning.

---

## *Resumé*

L'élevage de précision est un système qui introduit une dimension de bien-être animal tout en améliorant la gestion des troupeaux grâce aux avancées technologiques. Les stratégies de surveillance manuelle, couramment utilisées en élevage, présentent des contraintes de temps et de coût. Cette lacune peut permettre à des problèmes de santé de se développer jusqu'à devenir de véritables épidémies au sein des troupeaux, provoquant ainsi des perturbations des activités économiques et des menaces pour la santé publique, notamment en cas de zoonoses. L'élevage de précision innovant actuellement sur le marché, développé dans ce cadre, utilise les technologies d'intelligence artificielle, les caméras intelligentes et les capteurs IoT pour offrir une solution de surveillance continue de l'état de santé des animaux, fournir des alertes précoces en cas de maladies et permettre d'apporter un soutien approprié à chaque animal individuellement. D'un point de vue économique, cet assistant peut compenser certaines pertes en réduisant la transmission des maladies, entraînant ainsi une baisse des dépenses liées aux traitements vétérinaires et une diminution du nombre de décès. Les objectifs de cet assistant visent à permettre un système interactif et réactif face aux problèmes de santé et de comportement, afin de limiter les interventions humaines et d'accélérer les réponses sanitaires. Les outils sélectionnés incluent : le suivi individuel des animaux grâce à des caméras intelligentes, l'utilisation de capteurs IoT pour collecter des données sur l'environnement et le comportement, ainsi que l'utilisation d'algorithmes d'apprentissage profond (deep learning) pour l'analyse prédictive. Les actions clés seront soutenues par une plateforme de gestion des données, tandis que des systèmes de notification permettront des réponses rapides.

**Mots-clés :** IA, IoT, agriculture intelligente, surveillance, apprentissage profond.

## ملخص

يُعد الإنتاج الحيواني الدقيق نظامًا يُضيف بُعدًا خاصًا برعاية الحيوان، مع تحسين إدارة القطعان من خلال التقدّم التكنولوجي. تُعد استراتيجيات المراقبة اليدوية ممارسة شائعة في تربية الحيوانات، إلا أنها تعاني من قيود تتعلق بالوقت والتكلفة. ويمكن أن تؤدي هذه الفجوة إلى تطوّر التحديات الصحية لتصبح أوبئة خطيرة داخل القطعان، مما يُحدث اضطرابات في الأنشطة الاقتصادية ويُشكّل تهديدًا على الصحة العامة، لا سيما في حالة الأمراض المشتركة بين الإنسان والحيوان (الزoonoz).

يوفر نظام الإنتاج الحيواني الدقيق المبتكر، المتوفر حاليًا في السوق والذي تم تطويره في هذا الإطار، حلًا لمراقبة الحالة الصحية للحيوانات بشكل مستمر، من خلال استخدام تقنيات الذكاء الاصطناعي، الكاميرات الذكية، وأجهزة الاستشعار. كما يتيح النظام إطلاق إنذارات مبكرة في حالة ظهور أمراض، ويسمح بتقديم الدعم المناسب لكل حيوان على حدة (IoT).

من الناحية الاقتصادية، يمكن لهذا المساعد أن يُعوّض بعض الخسائر من خلال تقليل انتقال الأمراض، مما يؤدي إلى انخفاض النفقات المرتبطة بالعلاج البيطري، بالإضافة إلى تقليل عدد الوفيات.

تركّز أهداف هذا المساعد على تمكين نظام تفاعلي وذاتي الاستجابة للمشاكل الصحية والسلوكية، من أجل تقليل التدخلات البشرية، وتسريع الاستجابة من منظور صحي.

لجمع البيانات IoT تشمل الأدوات المختارة: التتبع الفردي للحيوانات باستخدام الكاميرات الذكية، استخدام أجهزة استشعار البيئية والسلوكية، واستخدام خوارزميات التعلم العميق لأغراض التحليل التنبؤي. وسيتم دعم الإجراءات الرئيسية من خلال منصة لإدارة البيانات، إلى جانب أنظمة إشعارات تُمكن من الاستجابات السريعة.

**الكلمات المفتاحية:** الذكاء الاصطناعي، إنترنت الأشياء، الزراعة الذكية، المراقبة، التعلم العميق

# Contents

|                                                                                 |             |
|---------------------------------------------------------------------------------|-------------|
| <b>Abstract</b>                                                                 | <b>i</b>    |
| <b>Résumé</b>                                                                   | <b>ii</b>   |
| <b>Contents</b>                                                                 | <b>iv</b>   |
| <b>List of Tables</b>                                                           | <b>vii</b>  |
| <b>List of Figures</b>                                                          | <b>viii</b> |
| <b>General Introduction</b>                                                     | <b>1</b>    |
| <b>1 Literature Review</b>                                                      | <b>2</b>    |
| 1.1 Introduction . . . . .                                                      | 2           |
| 1.2 Overview of Livestock Management . . . . .                                  | 2           |
| 1.3 Algeria's Sheep Crisis . . . . .                                            | 3           |
| 1.4 Diseases in Sheep and Their Indicators . . . . .                            | 4           |
| 1.5 Internet of Things . . . . .                                                | 6           |
| 1.6 Artificial Intelligence . . . . .                                           | 7           |
| 1.7 Machine learning . . . . .                                                  | 7           |
| 1.7.1 Definition . . . . .                                                      | 7           |
| 1.7.2 Methods . . . . .                                                         | 8           |
| 1.8 ML methods in agriculture: related work . . . . .                           | 12          |
| 1.9 Deep Learning . . . . .                                                     | 13          |
| 1.9.1 Convolutional neural networks (CNN) . . . . .                             | 13          |
| 1.9.2 Faster Region-Based Convolutional Neural Network (Faster R-CNN) . . . . . | 14          |
| 1.9.3 Recurrent Neural Network(RNN) . . . . .                                   | 15          |
| 1.10 DL methods in agriculture: related work . . . . .                          | 15          |
| 1.11 Current Market Solutions and Technologies . . . . .                        | 16          |
| 1.11.1 SmartShepherd Collar . . . . .                                           | 16          |
| 1.11.2 NB-IoT Project for Connected Sheep . . . . .                             | 17          |
| 1.11.3 GPS Collars for Sheep . . . . .                                          | 17          |
| 1.11.4 Ceres Tag Collars for Sheep . . . . .                                    | 18          |
| 1.11.5 Limits of Current Sheep Health Systems . . . . .                         | 19          |
| 1.12 Gaps and open issues . . . . .                                             | 20          |
| 1.13 Conclusion . . . . .                                                       | 20          |
| <b>2 System Design, System Requirements, and Architecture</b>                   | <b>22</b>   |
| 2.1 Introduction . . . . .                                                      | 22          |

|          |                                                        |           |
|----------|--------------------------------------------------------|-----------|
| 2.2      | Functional and Non-Functional Requirements . . . . .   | 22        |
| 2.2.1    | Functional Requirements . . . . .                      | 22        |
| 2.2.2    | Non-Functional Requirements . . . . .                  | 23        |
| 2.3      | System Overview . . . . .                              | 24        |
| 2.3.1    | General objective of the system . . . . .              | 24        |
| 2.4      | Hardware Components . . . . .                          | 25        |
| 2.4.1    | LoRa SX1278 with ESP32 . . . . .                       | 25        |
| 2.4.2    | MAX30102 Oxygen and Heart Rate Sensor . . . . .        | 26        |
| 2.4.3    | Vibration Sensor SW-420 . . . . .                      | 26        |
| 2.4.4    | GPS Module Neo-6M V2 . . . . .                         | 26        |
| 2.4.5    | Temperature Sensor DS18B20 . . . . .                   | 27        |
| 2.4.6    | Rechargeable Battery . . . . .                         | 27        |
| 2.4.7    | 3D-printed cover for a belt . . . . .                  | 28        |
| 2.5      | Software Components . . . . .                          | 28        |
| 2.5.1    | Development environment . . . . .                      | 28        |
| 2.5.2    | Programming language . . . . .                         | 29        |
| 2.5.3    | Libraries used . . . . .                               | 29        |
| 2.6      | System Architecture . . . . .                          | 30        |
| 2.6.1    | Deployment on the Animal: . . . . .                    | 32        |
| 2.6.2    | Challenges and Technical Considerations: . . . . .     | 32        |
| 2.6.3    | Data Collection . . . . .                              | 32        |
| 2.6.4    | Data Storage and Transmission . . . . .                | 33        |
| 2.6.5    | Data Cleaning and Preprocessing Techniques . . . . .   | 34        |
| 2.6.6    | Labeling and Ground Truth Collection . . . . .         | 38        |
| 2.6.7    | AI Models . . . . .                                    | 38        |
| 2.7      | Model Training and Validation . . . . .                | 44        |
| 2.8      | Symptom-to-Sensor Mapping . . . . .                    | 44        |
| 2.9      | Conclusion . . . . .                                   | 45        |
| <b>3</b> | <b>Implementation, Evaluation and Results</b>          | <b>46</b> |
| 3.1      | Introduction . . . . .                                 | 46        |
| 3.2      | Experimental Setup and Scenarios . . . . .             | 47        |
| 3.2.1    | Experimental Setup . . . . .                           | 47        |
| 3.2.2    | Experimental Scenarios . . . . .                       | 50        |
| 3.3      | Models hyperparameters . . . . .                       | 51        |
| 3.3.1    | LSTM Model Parameters . . . . .                        | 52        |
| 3.3.2    | SVM Model hyperparameters . . . . .                    | 52        |
| 3.3.3    | XGBOOST Model hyperparameters . . . . .                | 52        |
| 3.3.4    | LightGBM Model hyperparameters . . . . .               | 52        |
| 3.3.5    | GRU Model hyperparameters . . . . .                    | 53        |
| 3.4      | Dataset . . . . .                                      | 53        |
| 3.5      | Evaluation metrics . . . . .                           | 54        |
| 3.6      | Quantitative Findings . . . . .                        | 55        |
| 3.7      | Comparative Analysis . . . . .                         | 58        |
| 3.8      | System Development and Deployment Challenges . . . . . | 59        |
| 3.9      | Ethical and Societal Implications . . . . .            | 59        |
| 3.10     | Conclusion . . . . .                                   | 60        |
|          | <b>General Conclusion</b>                              | <b>61</b> |

---

## CONTENTS

---

|                       |           |
|-----------------------|-----------|
| <b>Bibliography</b>   | <b>62</b> |
| <b>Webographie</b>    | <b>70</b> |
| <b>Start-up Annex</b> | <b>70</b> |

# List of Tables

|     |                                                                                                                         |    |
|-----|-------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | Summary of common sheep diseases and their key observable indicators for early detection and health monitoring. . . . . | 6  |
| 2.1 | List of sensors and their roles in health monitoring. . . . .                                                           | 33 |
| 2.2 | Transformation of Clinical Symptoms into Measurable Data. . . . .                                                       | 44 |
| 3.1 | Healthy Sheep Normal Sensor Values. . . . .                                                                             | 50 |
| 3.2 | Infected Sheep Abnormal Sensor Values (PPR case) . . . . .                                                              | 51 |
| 3.3 | Sheep in Alert. . . . .                                                                                                 | 51 |
| 3.4 | Comparison of model performance. . . . .                                                                                | 58 |
| 1.5 | Project Completion Schedule . . . . .                                                                                   | 73 |
| 1.6 | Prototype Development Cost Breakdown . . . . .                                                                          | 78 |

# List of Figures

|      |                                                                    |    |
|------|--------------------------------------------------------------------|----|
| 1.1  | Traditional breeding in Algeria [W1]                               | 3  |
| 1.2  | Number of sheep in Algeria drops from 36 to 17 million [W67][W68]  | 3  |
| 1.3  | Examples of internet of things applications [W2]                   | 6  |
| 1.4  | Machine Learning Methods [W4]                                      | 8  |
| 1.5  | Supervised learning Workflow [W5]                                  | 8  |
| 1.6  | Example in Random Forest [W8]                                      | 9  |
| 1.7  | Example in Support Vector Machine [W10]                            | 9  |
| 1.8  | Example in k Nearest Neighbour [W11]                               | 10 |
| 1.9  | Example in Decision tree [W13]                                     | 10 |
| 1.10 | Example in eXtreme Gradient Boosting [W17]                         | 11 |
| 1.11 | Unsupervised learning techniques.                                  | 11 |
| 1.12 | Reinforcement learning, agent and environment interactions.        | 12 |
| 1.13 | Deep Learning Model with Two Hidden Layers [W19].                  | 14 |
| 1.14 | CNN model                                                          | 14 |
| 1.15 | the Recurrent Neural Network model in deep learning.               | 15 |
| 1.16 | SmartShepherd Collar [W31]                                         | 17 |
| 1.17 | NB-IoT Project for Connected Sheep [W32]                           | 17 |
| 1.18 | GPS Collars for Sheep [W34]                                        | 18 |
| 1.19 | Ceres Tag Collars for Sheep [W35]                                  | 18 |
| 2.1  | Overview of how the smart collar works.                            | 25 |
| 2.2  | ESP32 LoRa Microcontroller Board [W36]                             | 25 |
| 2.3  | MAX30102 Pulse Oximeter and Heart Rate Monitor [W37]               | 26 |
| 2.4  | SW-420 Vibration Sensor Module [W38]                               | 26 |
| 2.5  | NEO-6M GPS Module [W39]                                            | 27 |
| 2.6  | Temperature Sensor DS18B20 [W40]                                   | 27 |
| 2.7  | Rechargeable Battery Module [W41]                                  | 28 |
| 2.8  | system architecture                                                | 31 |
| 2.9  | Overview of the IoT-Based Sheep Health Monitoring Process.         | 32 |
| 2.10 | Field Setup During Dataset Collection for the Smart Collar System. | 33 |
| 2.11 | Converting data types.                                             | 34 |
| 2.12 | Converting data types.                                             | 35 |
| 2.13 | Overview of Common Data Normalization Techniques in Preprocessing. | 35 |
| 2.14 | Outlier Detection and Removal Using the IQR Method.                | 37 |
| 2.15 | Outlier Detection and Removal Using the IQR Method.                | 38 |
| 2.16 | Data Labeling for Smart Collar.                                    | 39 |
| 2.17 | Example of Labeled Dataset for Smart Collar System.                | 39 |
| 2.18 | The architecture of the LSTM method [W42]                          | 40 |
| 2.19 | The architecture of the XGBOOST method [43]                        | 40 |



|      |                                                                      |    |
|------|----------------------------------------------------------------------|----|
| 2.20 | The architecture of the SVM method [44]                              | 41 |
| 2.21 | The architecture of the LightGBM method [45]                         | 42 |
| 2.22 | The architecture of the GRU method [46]                              | 43 |
|      |                                                                      |    |
| 3.1  | ESP32-Based Health Monitoring System Circuit Simulation              | 47 |
| 3.2  | Smart Collar Prototype for Early Detection of PPR in Small Ruminants | 48 |
| 3.3  | Field Test of the Smart Collar on a Sheep                            | 49 |
| 3.4  | The confusion matrix for LSTM model.                                 | 55 |
| 3.5  | The evaluation metrics plots for LSTM model.                         | 55 |
| 3.6  | The confusion matrix for SVM model.                                  | 56 |
| 3.7  | The evaluation metrics plots for SVM model.                          | 56 |
| 3.8  | The confusion matrix for XGBoost model.                              | 56 |
| 3.9  | The evaluation metrics plots for XGBoost model.                      | 56 |
| 3.10 | The confusion matrix for GRU model.                                  | 57 |
| 3.11 | The evaluation metrics plots for GRU model.                          | 57 |
| 3.12 | The confusion matrix for LightGBM model.                             | 57 |
| 3.13 | The evaluation metrics plots for LightGBM model.                     | 57 |
| 3.14 | The LSTM Model's ROC Curve and AUC Score.                            | 58 |

# General Introduction

In recent years, the health management of livestock, especially sheep, has become a growing concern due to the economic and epidemiological impacts of diseases such as "*peste des petits ruminants (PPR)*". Traditional disease detection methods are largely reactive, relying on visual inspections and delayed clinical interventions. With the advent of the Internet of Things (IoT) and Artificial Intelligence (AI), there is a compelling opportunity to transform animal health monitoring through intelligent, automated systems. This project is motivated by the need for early, accurate, and non-invasive detection of PPR using a smart collar equipped with appropriate sensors that can continuously monitor vital signs and behavior.

The primary goal is to design and implement an integrated system that collects sensor data, preprocesses it, and uses AI models to classify sheep as healthy or infected. This study is guided by the following hypotheses: – a smart collar equipped with appropriate sensors can detect early symptoms of PPR with high accuracy; – AI models trained on multimodal sensor data will outperform traditional threshold-based systems in disease detection; – the use of such a system will reduce the time between infection and intervention in a flock. To achieve this, the project will address several research questions regarding sensor relevance, data processing techniques, and model effectiveness.

This thesis is organized into four chapters:

- Chapter 1 provides a literature review on livestock management practices, common sheep diseases and indicators, enabling technologies like IoT and AI, existing solutions, and current gaps.
- Chapter 2 presents the system design and architecture, including requirements, hardware and software components, sensor deployment, data handling, and model development.
- Chapter 3 details the implementation and evaluation of the system, discusses performance results from real-world testing, and reflects on challenges, limitations, and ethical considerations.
- Startup annex presents a comprehensive overview of the business plan and market strategies.

# Literature Review

## 1.1 Introduction

The rise of Internet of Things (IoT) technologies and artificial intelligence is profoundly transforming agricultural practices, particularly in the field of animal health monitoring. However, existing systems remain limited in their ability to detect certain specific diseases, notably "*peste des petits ruminants (PPR)*", a highly contagious viral disease that poses a major threat to sheep farming. In this context, this thesis proposes the development of a smart collar dedicated to the early detection of PPR in sheep. Based on physiological sensors (temperature, heart rate, oxygen level, activity), an IoT architecture, and embedded data processing algorithms, the device aims to identify characteristic clinical signs of the disease before visible symptoms appear. This project builds upon recent work on the structuring and classification of agricultural sensor data, which emphasizes the importance of organizing IoT data according to specific applications for better reuse and targeted detection [LBHH23]. The objective is to design a system that is practical, scalable, and accessible, suitable for field conditions and capable of enhancing farm resilience against epidemics.

## 1.2 Overview of Livestock Management

Sheep breeding is vital to Algeria's economy since it plays a significant role in the production of red meat and the livestock industry. The native breed known as "Rumbi," which is renowned for its hardiness and productivity, is especially notable in the Tiaret region. This rectangular breed accounts for approximately 11% of Algeria's total sheep population, with rams reaching up to 80 kg. With 29 million sheep and an annual production of over 5,291,695 quintals of meat, Algeria's sheep industry is a major player in the Maghreb region. Sheep meat continues to be the most popular in the area, despite competition from beef and chicken. It has a major economic impact, contributing over

half of the country's red meat production and making up 10 to 15% of GDP, or roughly \$5 millions. Additionally, it accounts for 35% of all agricultural production[ZHZ<sup>+</sup>24].



Figure 1.1 – Traditional breeding in Algeria [W1]

### 1.3 Algeria's Sheep Crisis

Algeria's Sheep population is worse Than ever. As a result of infectious diseases Sheep herds in Algeria have been progressively dropping in recent years, despite a notable increase in cattle prices. The spread of harmful and contagious diseases that are affecting an increasing number of farms is mostly to blame for this predicament. Many farmers are compelled to close their businesses or substantially cut back on their cow herds as a result of these health hazards. Innovative technical solutions must be put into place immediately in order to maintain this crucial industry. Limiting losses, protecting flocks, and maintaining market stability can be achieved with the use of preventive measures, ongoing monitoring systems, and early detection technology.

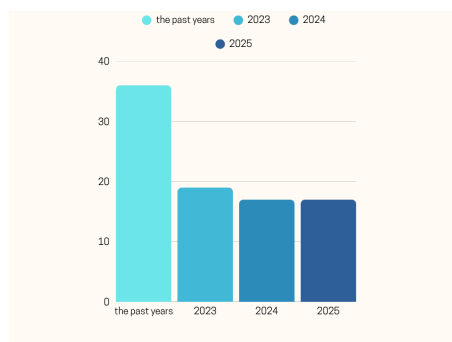


Figure 1.2 – Number of sheep in Algeria drops from 36 to 17 million [W67][W68]

## 1.4 Diseases in Sheep and Their Indicators

| Diseases                                    | Indicators                                                                                                                              |
|---------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| Infections parasitaires gastro-intestinales | Weight loss, dehydration (flabby skin), staying away from the group, lying down frequently                                              |
| Lameness                                    | Abnormal gait, avoiding the position of the affected limb, often standing on three legs, prolonged periods.                             |
| Ectoparasite infections                     | Excessive scratching, areas of loss of wool, restless behaviors, such as rubbing against objects or biting themselves                   |
| Mastitis                                    | Swollen udder, red or warm, signs of pain when approached or trying to feed, lethargic behavior and reluctance to move                  |
| Stress                                      | Incessant movements, frequent bleating, or attempts to escape. They may also stay away from others                                      |
| Respiratory Diseases                        | Cough, tense neck, rapid breathing                                                                                                      |
| Bluetongue disease                          | Swelling of the head and tongue, nasal discharge, skin lesions, high fever, respiratory distress                                        |
| Scrapie disease                             | Behavioral changes, itching, loss of coordination, and abnormal movements. Sheep may rub against objects and show signs of nervousness. |

| <b>Diseases</b>                  | <b>Indicators</b>                                                                                                                                                                                             |
|----------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Dermatophilosis (skin infection) | Bursts, sores on the skin, itching, inflammation, areas of thinning coat                                                                                                                                      |
| Malnutrition                     | Visible weight loss, dull coat, and weakness                                                                                                                                                                  |
| Digestive Problems               | Abnormal posture (such as lying on the stomach), lack of appetite                                                                                                                                             |
| Tension                          | Agitated behavior, rapid movements, frequent vocalizations, rigid posture                                                                                                                                     |
| Paratuberculosis                 | Weight loss, Chronic diarrhea, Weakness, Reduction in food intake, Dehydration                                                                                                                                |
| Anthrax                          | High body temperature, Respiratory distress, cough and wheezing, Swelling and red-dening of the skin.                                                                                                         |
| Dermatophytosis                  | Dry, scaly skin crusts, Hair loss, Pruritus, Inflammation and redness of the skin, Touch sensitivity, Lameness, and difficulty walking.                                                                       |
| Pasteurella                      | Persistent cough, Respiratory distress, Mucopurulent nasal discharge, Fever, Loss of appetite, Generalized weakness.                                                                                          |
| Enterotoxemia                    | animal in shock , abdominal pain, high fever, diarrhea with pieces of mucous membrane and blood, dehydration, nervous signs (agitation, opisthotonos (tense animal, with head thrown back) and convulsions) . |

| Diseases      | Indicators                                                                                                                                        |
|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------|
| Toxoplasmosis | Spontaneous abortions, Birth of weak lambs, Increased body temperature, Reduced appetite, Diarrhea.                                               |
| Peste (PPR)   | High fever, Dejection and weakness, Cough and difficulty breathing, Loss of appetite, Nasal / eye discharge, Mundus lesions, Abortions, Diarrhea. |

Table 1.1 – Summary of common sheep diseases and their key observable indicators for early detection and health monitoring.

## 1.5 Internet of Things

The Internet of Things (IoT) refers to a dynamic global network where physical and virtual objects equipped with unique identities and intelligent interfaces are interconnected and integrated into the digital environment. These “smart things” can communicate, collect and exchange data, respond autonomously to real-world events, and trigger actions or services, often without human intervention. Interfaces allow users to query, monitor, and control these devices remotely, while ensuring privacy and security. The rapid growth of IoT has made it one of the largest sources of big data, characterized by high volume, velocity, variety, and spatio-temporality [KBF<sup>+</sup>22]. This evolution brings critical challenges in real-time indexing, processing, and managing IoT data for applications such as agriculture, healthcare, and smart cities.

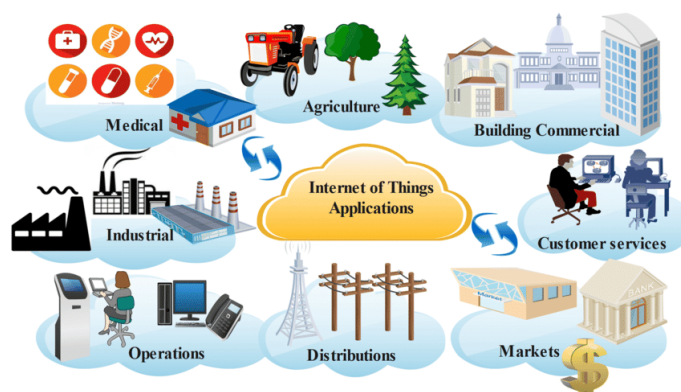


Figure 1.3 – Examples of internet of things applications [W2]

## 1.6 Artificial Intelligence

The manifestations of intelligence typically involve comprehending a situation and effectively solving difficulties, especially when the problem-solving skill is adequate to provide multiple answers for the given problem. Intelligence is also evident in the decision-making process, where the most suitable solution for the present problem is chosen [Ber20].

While the machine has the capability to solve difficulties, it remains uncertain whether it can acquire the ability to comprehend a specific problem and make the appropriate decision to solve it. How closely can machine capabilities approximate those of human beings? In 1956, a symposium held at Dartmouth University delved into these inquiries and resulted in the emergence of the word 'AI' [Jam06], which stands for Artificial Intelligence.

Artificial intelligence focuses on the study of intelligent behavior in machines, encompassing perception, reasoning, learning, communication, and action in complex environments. Therefore, AI has two main objectives: to develop machines that can perform these tasks as well as or better than humans, and to understand this behavior whether it is exhibited by machines or humans [Nil98].

In addition to these goals, the conclusion, that rather than teaching computers everything they need to know how to carry out jobs. They might be able to learn on their own, and the recent development of the internet was one of the most significant breakthroughs that caused machine learning to arise and accelerate the advancement of artificial intelligence.

## 1.7 Machine learning

### 1.7.1 Definition

Learning is the foundational methodology of most machine learning (ML) approaches. Its goal is to acquire the necessary experience (training data) to successfully complete a task. A set of features, often referred to as attributes or variables, is generally used to characterize a single instance. Nominal features are those that are listed: binary features take values of 0 or 1, ordinal features take values such as A or B, and numerical features take specific values ranging from 0 to 1 (integers, real numbers, etc.). To evaluate the performance of an ML model for a given task, one can look at its performance metrics, which tend to improve as more data is processed. The efficiency of ML models and algorithms can be estimated using various statistical and mathematical approaches. Once the training phase is complete, the learned model can be used to make predictions or cluster new data (test data) based on what it learned during the training phase. Several statistical and mathematical models can be used to estimate the performance of an ML model or algorithm. After the learning phase, the trained model can be used to make predictions or group new data (test data) based on the knowledge acquired during the training process.



Figure 1.4 illustrates a common machine learning strategy[MKH<sup>+</sup>24].

## 1.7.2 Methods

As shown in the graph below, machine learning is generally classified into three main categories: supervised learning, unsupervised learning, and reinforcement learning. Each of these categories has specific goals and approaches.

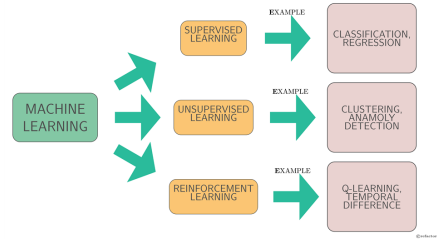


Figure 1.4 – Machine Learning Methods [W4]

### a) Supervised Learning

The machine learning task of supervised learning involves using sample input-output pairs to train a function that maps an input to an output. From labelled training data, which consists of a collection of training instances, it infers a function. The machine learning algorithms that require external aid are known as supervised algorithms. The train and test datasets are separated from the input dataset. An output variable from the train dataset needs to be categorized or forecasted. Every algorithm picks up some sort of pattern from the training dataset and uses it to predict or classify data from the test dataset [Mah20]. There are many categories

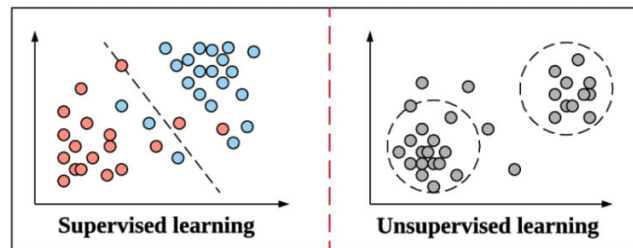


Figure 1.5 – Supervised learning Workflow [W5]

of algorithms for supervised learning [W6] : A supervised classification approach that can be applied to regression and classification issues is the Random Forest Classifier [Pav00]. This classifier is referred to as an overall classifier because it uses multiple classification algorithms either the same or different to categorize things. Since each decision tree is a single classifier and the objective prediction is

based on the majority voting method, it involves, as its name implies, establishing a forest of decision trees, the more trees in the forest, the more exact the findings [Bre94]. Support vector machines (SVMs) are a supervised machine learning

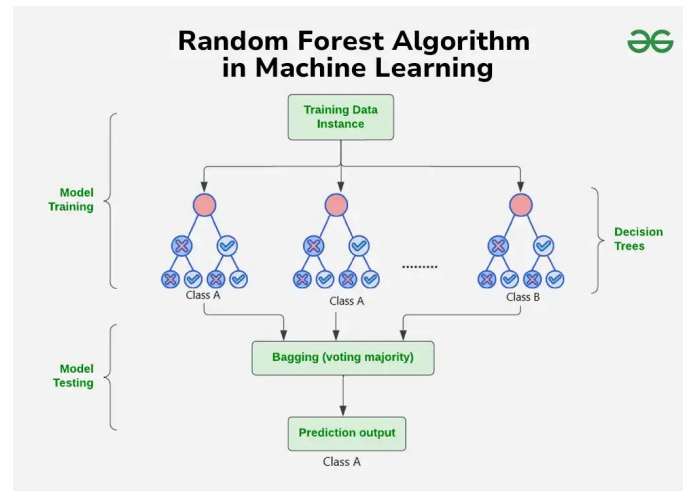


Figure 1.6 – Example in Random Forest [W8]

algorithm used for classification and regression. They are based on statistical learning theory and aim to find an optimal hyperplane that separates data in a way that maximizes the margin between classes. To handle nonlinear problems, SVMs use kernel functions to project the data into a higher-dimensional space where it becomes linearly separable. This algorithm is particularly effective for small or medium-sized datasets and in cases where a clear separation between classes is required[MCLL24]. The k-Nearest Neighbors (kNN) technique is a supervised

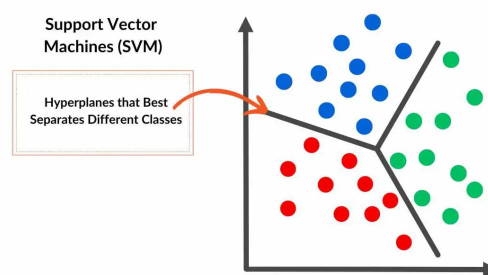


Figure 1.7 – Example in Support Vector Machine [W10]

learning approach that finds widespread application in different sectors, including data mining, recommendation systems, and the Internet of Things (IoT), contributing significantly to the advent of Industry 4.0. The k-Nearest Neighbors (kNN) method is an instance-based nonparametric learning algorithm, often used for supervised tasks such as classification and regression. It makes its predictions by scanning the composition of live data as new instances are introduced, without the need for

a separate training step beforehand[HUU<sup>+</sup>24]. The decision tree is defined as a

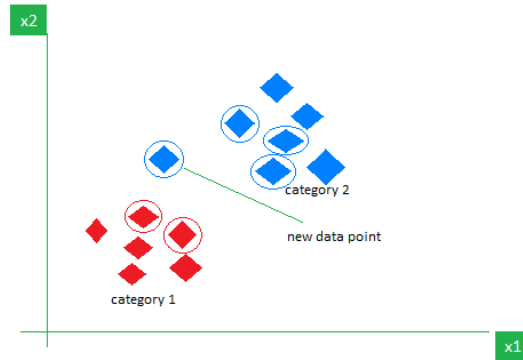


Figure 1.8 – Example in k Nearest Neighbour [W11]

supervised learning model that hierarchically maps a data domain onto a response set. It divides a data domain (node) recursively into two subdomains such that the subdomains have a higher information gain than the node that was split. We know the goal of supervised learning is the classification of the data, and therefore, the information gain means the ease of classification in the subdomains created by a split. Finding the best split that gives the maximum information gain (i.e., the ease of classification) is the goal of the optimization algorithm in the decision tree-based supervised learning [SS16, Mah20] XGBoost is a distributed, open-source machine

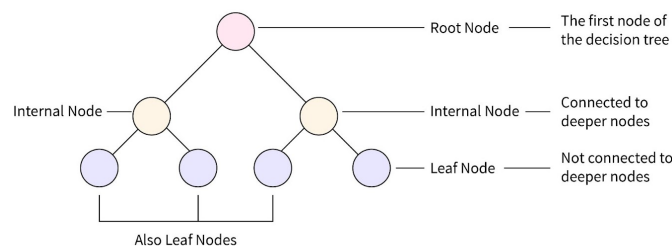


Figure 1.9 – Example in Decision tree [W13]

learning library that uses gradient-boosted decision trees, a supervised learning boosting algorithm using gradient descent. It is recognized for its speed, efficiency, and ability to scale to large datasets [MCLL24].

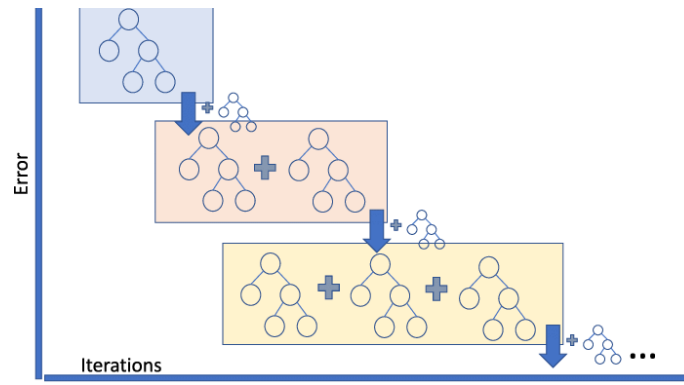


Figure 1.10 – Example in eXtreme Gradient Boosting [W17]

**Regression :** A supervised learning method called regression is used to forecast continuous responses[FKL<sup>+</sup>13]. It is one of the most significant and widely used machine learning and statistics techniques; it enables the creation of predictions from data by figuring out how elements in the data relate to the output, which is continuous and observed. There are other models that can be employed, including Random Forest, Polynomial, Support Vector [AK15], and Decision Tree regressions. The most basic model is the Simple Linear regression[Wei05].

#### b) Unsupervised Learning :

Unsupervised Learning Techniques: with these methods, the learning algorithm is left to identify similarities between its input data because the training data is not labeled. It makes an effort to group the supplied dataset into clusters or classes. The input dataset consists of unprocessed data without any target results or class labels. Unstructured datasets lacking class labels, optimization criteria, and feedback are referred to as raw data [SMS20]. Unsupervised learning aims to detect natural

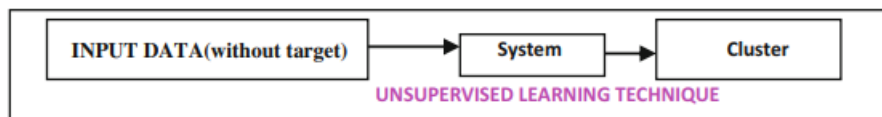


Figure 1.11 – Unsupervised learning techniques.

divisions in datasets, benefiting preprocessing and feature extraction systems. A key technique is clustering, which groups similar data points based on proximity (e.g., using Euclidean or cosine distance), differing from supervised learning by working with unlabeled data. Popular clustering algorithms include K-means, Mean-Shift, Hierarchical Clustering, and K-NN.

### c) Reinforcement Learning :

It is a method of machine machine learning that depends on rewarding desired behaviors or punishing undesirable behaviors. In this way, it makes the robot capable of making appropriate decisions or actions by taking advantage of its awareness of the environment surrounding it. It is based on the principle of trial and error. three three major composite make up the reinforcement learning :

**Agent:** the learner and the decision maker .

**Environment:** where the agent learns and decides what actions to perform.

**Action:** a set of actions which the agent can perform .

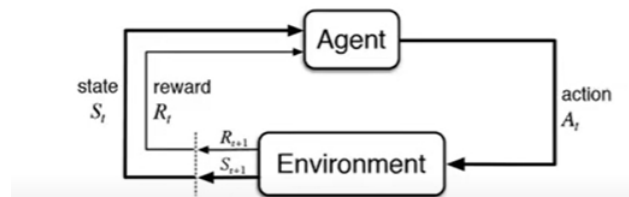


Figure 1.12 – Reinforcement learning, agent and environment interactions.

The goal is defined optimal policy in the sense of maximizing the expected value of the total reward over any no successive steps starting from the current state . One of the reinforcement learning algorithm is Q-learning algorithm .

## 1.8 ML methods in agriculture: related work

Several studies have explored the use of machine learning methods in agriculture and animal health to identify and prevent diseases in livestock. This section provides a comparative analysis of major studies using machine learning models in various sheep-related contexts.

A study [ZZW<sup>+</sup>18] classified sheep grazing and rumination behaviors using supervised algorithms (Random Forest, SVM, k-NN, AdaBoost) on 27,317 sensor data points with 39 features. Random Forest achieved the highest accuracy (92%), showing the value of proper feature selection for model performance. Another study [KVK<sup>+</sup>24] highlighted the influence of management and environment on health by estimating subclinical mastitis in dairy ewes using machine learning (clustering and supervised models), reaching 96.3% accuracy with SVM. In [MCLL24], the authors used machine learning and mobile sensors to predict travel stress in sheep, with XGBoost attaining 94.92% accuracy. Although only 195 sheep were used for testing, the method allows for real-time stress assessment. Another study [HMB<sup>+</sup>24] used IMU sensors on goats and sheep to analyze behavior from over 13 million data

samples. Similar movements made some actions hard to distinguish, but advanced preprocessing (L2 norm, DFT) improved accuracy. They noted limits due to small sample size and suggested adding more sensors. The work done in Another study [ALECRM<sup>+</sup>23] reviewed 263 studies on sensor fusion for animal monitoring, finding most focus on cows and horses, with only 8.7% on sheep. Common sensors are accelerometers, gyroscopes, and cameras, mainly in collars and ear tags. They noted a lack of sensors for key health parameters like heart and respiratory rates. Noor [NKA<sup>+</sup>23] used Random Forest and SVM to detect abnormal sheep behaviors like lameness and feeding problems, with RF reaching 82.5% accuracy. These findings show that ML models work well for behavior analysis but need additional physiological data (e.g., heart rate, temperature) to boost accuracy and support early disease detection. Turner [TSH<sup>+</sup>23] used accelerometer data and machine learning (RF, SVM) to detect lambing behaviors with up to 87.6% accuracy. Despite this progress, results are affected by environmental factors and sheep variability. Goncalves [GdRMB<sup>+</sup>24] used inertial sensors and ML models (RF, SVM, k-NN) to detect parturition in goats, identifying behavioral changes up to 4 hours before labor. RF predicted timing with 61% accuracy, but performance was limited by a small dataset and change detection challenges.

## 1.9 Deep Learning

By using a hierarchy of concepts, deep learning is a type of machine learning that lets computers learn from experience and comprehend the world. The computer taught itself how to handle and absorb data [Kim16]. It operates on a network of synthetic neurons that are modeled after the human brain. The neurons in this network are arranged in tens or even hundreds of layers, each of which receives and processes data from the layer above it.

Neural networks and deep learning differ in the depth of the model; deep learning is a term for intricate neural networks. It has to do with feature extraction and transformation, which try to establish a connection between inputs and the corresponding neural responses that the brain produces. Elaborate patterns of how information can flow across the model confer complexity.

### 1.9.1 Convolutional neural networks (CNN)

Computer vision and image classification applications primarily rely on convolutional neural networks (CNNs or ConvNets). These networks are capable of identifying features and patterns in images and videos, enabling tasks such as object localization, image identification, shape differentiation, and facial recognition.

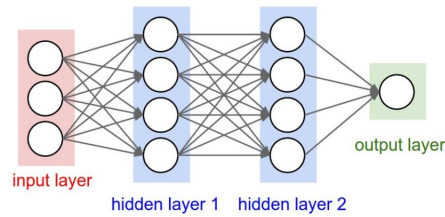


Figure 1.13 – Deep Learning Model with Two Hidden Layers [W19].

CNNs utilize principles of linear algebra, particularly matrix multiplication, to detect patterns within an image. CNNs are a specific type of neural network composed of multiple layers of nodes. They include an input layer, one or more hidden layers, and an output layer. Each node is connected to another and is associated with a weight and a threshold. If the output of a given node exceeds the defined threshold, it activates and passes the information to the next stage in the network. Otherwise, no information is transmitted to the upper layer[HMB<sup>+</sup>24].

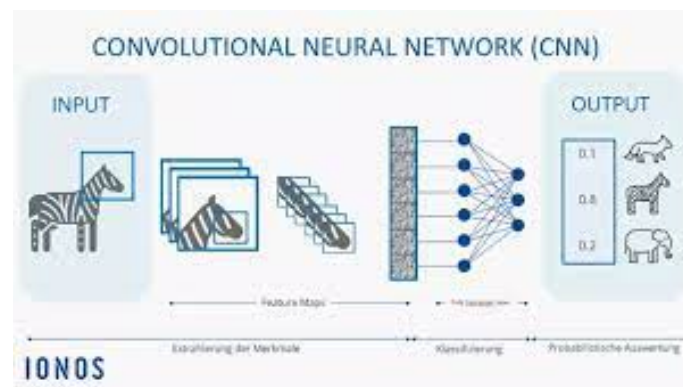


Figure 1.14 – CNN model

### 1.9.2 Faster Region-Based Convolutional Neural Network (Faster R-CNN)

Faster R-CNN is an object detection model that identifies objects in an image, draws bounding boxes around them, and determines their classification. It operates through a two-stage process:

Step 1: The Region Proposal Network (RPN) generates candidate regions that may contain objects.

Step 2: These regions are then analyzed to predict the object category and refine the bounding box to more precisely match the identified object[HMB<sup>+</sup>24].

### 1.9.3 Recurrent Neural Network(RNN)

A Recurrent Neural Network, or RNN, is a deep learning model designed to be trained on sequential data or time series, allowing the machine learning (ML) algorithm to make predictions based on data sequences or to draw conclusions from sequential inputs. For example, an RNN could be used to forecast daily flood levels based on weather information, tides, and past flooding data. However, these networks can also be applied to ordinal or temporal problems such as language translation, natural language processing (NLP), sentiment analysis, speech recognition, and image captioning. The most popular variants of recurrent neural network architectures include: Standard RNN, Bidirectional Recurrent Neural Networks (BRNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Encoder–Decoder RNNs [W26]. A comprehensive description of the architecture will be provided in the second chapter.

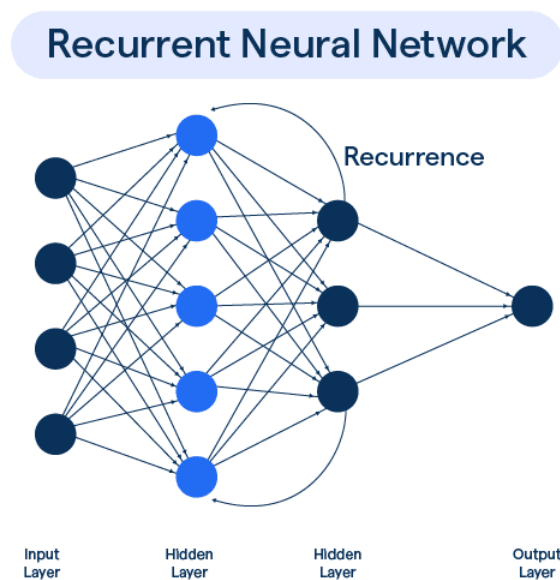


Figure 1.15 – the Recurrent Neural Network model in deep learning.

## 1.10 DL methods in agriculture: related work

Several studies have explored the use of deep learning methods in agriculture and animal health to identify and prevent diseases in livestock. This section provides a comparative analysis of major studies using deep learning models in various sheep-related contexts.



Turner [TSH<sup>+</sup>23] used accelerometer data with a CNN-LSTM model to detect lambing behaviors in sheep, achieving 87.6% accuracy, but results are limited by small data size and individual variability. Rohan [RRH<sup>+</sup>24] reviewed 23 deep learning models for farm animal behavior detection, finding YOLOv5 and Faster R-CNN best for image recognition, and ResNet50/VGG16 effective for feature extraction. They noted the lack of standardized data collection limits cross-study comparisons. In [TTH<sup>+</sup>23], the authors improved sheep behavior classification using LSTM, BLSTM, and Random Forest with SMOTE for data imbalance. LSTM reached 88% accuracy and BLSTM an F1-score of 0.84, but RF performed worse, and individual variations still limit robustness. Noor [NKA<sup>+</sup>23] combined GPS, accelerometers, and cameras to monitor sheep with ML. YOLOv5 (95–98%), VGG16 (94%), VGG17 (96%), and LSTM (87–90%) showed high accuracy. Multi-sensor fusion improved classification but results depend on environment and data quality. Works done by Hollevoet [HMB<sup>+</sup>24] used accelerometers and ML models (MLP, CNN, hCNN) to analyze goat behavior. hCNN and MLP reached 94% accuracy, CNN 89.8%, but distinguishing similar actions was difficult and the small sample size limits generalization. Arshad [ABV<sup>+</sup>24] used smart sensors and deep learning (YOLOv8n-CBAM, CNN) to monitor sheep behavior, achieving 97.7% (head detection) and 93% (behavior classification). Despite high accuracy, challenges include data standardization and limited access for small farmers.

## 1.11 Current Market Solutions and Technologies

Here is a comparison of the different sheep behavior detecting technologies currently in use. Every system is thoroughly explained, emphasizing its features and requirements, and is, if feasible, illustrated.

### 1.11.1 SmartShepherd Collar

**Year:** 2018 **Country:** New Zealand

**Location:** Around the neck of sheep (ewes and lambs)

**What it detects:**

- **Proximity Data:** Tracks the distance between ewes and their lambs to monitor interactions.
- **Interaction Frequency:** Records the frequency of interactions between ewes and lambs, providing insights into maternal bonding.
- **Pedigree Information:** Links animal identification tags to track lineage and rearing ranks of lambs.

**Battery Duration:** Not specified.[30]



Figure 1.16 – SmartShepherd Collar [W31]

### 1.11.2 NB-IoT Project for Connected Sheep

**Year:** 2020

**Country:** Norway

**Location:** Around the necks of sheep (in the form of connected collars)

**What it detects:** Real-time location tracking and behavior/welfare data collection.

**Battery Duration:** Not specified.

[ALC<sup>+</sup>24]



Figure 1.17 – NB-IoT Project for Connected Sheep [W32]

### 1.11.3 GPS Collars for Sheep

**Year:** Not specified

**Country:** France

**Location:** Around the neck of the sheep

**What it detects:** Provides real-time location, records movements, and ensures continuous monitoring of sheep position.

**Battery Duration:** Approximately one year under normal conditions.[W33]



Figure 1.18 – GPS Collars for Sheep [W34]

#### 1.11.4 Ceres Tag Collars for Sheep

**Year:** 2020 (Commercial Launch) **Country:** Australia/New Zealand (Global Use)

**Location:**

- Collar-mounted (for adult sheep)
- Ear tag option (RFID + sensors for lambs)

**What it detects:**

- **Interaction Patterns:** Infers bonding behavior through movement algorithms and cluster detection
- **Pedigree Tracking:** Integrates with national RFID systems (NLIS in AU/NZ) for lineage documentation
- **Health Indicators:** Detects abnormal activity levels signaling illness or distress
- **Predation Alerts:** Identifies sudden movement changes indicating predator attacks

**Battery Duration:** 5 years (non-replaceable lithium battery). [W35]



Figure 1.19 – Ceres Tag Collars for Sheep [W35]

### 1.11.5 Limits of Current Sheep Health Systems

The table indicates the need for more comprehensive and disease-specific solutions by demonstrating that while existing smart systems for sheep health monitoring (such as NB-IoT and GPS collars) can identify certain problems, they fall short in addressing numerous ailments.

| Disease                                                                             | SmartShepherd | NB-IoT | GPS Collars | Ceres Tag Collars |
|-------------------------------------------------------------------------------------|---------------|--------|-------------|-------------------|
| Gastrointestinal Parasitic Infections [B <sup>+</sup> 23, A <sup>+</sup> 21b, KG09] | No            | No     | Yes         | No                |
| Lameness [RGBR20, A <sup>+</sup> 21b, G <sup>+</sup> 20]                            | No            | Yes    | Yes         | No                |
| Ectoparasite Infections [ZZZ22, A <sup>+</sup> 21b]                                 | No            | No     | Yes         | No                |
| Mastitis [RGBR20, A <sup>+</sup> 21b, R <sup>+</sup> 07]                            | Yes           | Yes    | Yes         | Yes               |
| Stress [A <sup>+</sup> 21b, G <sup>+</sup> 20, ER <sup>+</sup> 20]                  | No            | Yes    | Yes         | No                |
| Respiratory Diseases [A <sup>+</sup> 21b, AER <sup>+</sup> 18, A <sup>+</sup> 21a]  | No            | Yes    | Yes         | No                |
| Bluetongue Disease [M <sup>+</sup> 14, A <sup>+</sup> 18]                           | No            | No     | No          | No                |
| Scrapie [M <sup>+</sup> 14, R <sup>+</sup> 07]                                      | No            | No     | No          | No                |
| Dermatophilosis (Skin Infection) [A <sup>+</sup> 18]                                | No            | No     | No          | No                |
| Malnutrition [A <sup>+</sup> 21b, AEM <sup>+</sup> 20, KG09]                        | No            | Yes    | Yes         | No                |
| Digestive Problems [RGBR20, A <sup>+</sup> 21b, KG09]                               | No            | Yes    | Yes         | No                |
| Tension [A <sup>+</sup> 21b, G <sup>+</sup> 20, ER <sup>+</sup> 20]                 | No            | Yes    | Yes         | No                |
| Paratuberculosis [A <sup>+</sup> 21b, R <sup>+</sup> 07]                            | No            | No     | No          | No                |
| Anthrax [AER <sup>+</sup> 18, OIE22]                                                | No            | No     | No          | No                |
| Dermatophytosis [A <sup>+</sup> 18]                                                 | No            | No     | No          | No                |
| Pasteurella [A <sup>+</sup> 21b, R <sup>+</sup> 18]                                 | No            | No     | No          | No                |
| Enterotoxemia [A <sup>+</sup> 21b, R <sup>+</sup> 07]                               | No            | No     | No          | No                |

|                                                                            |    |    |    |    |
|----------------------------------------------------------------------------|----|----|----|----|
| Toxoplasmosis<br>[R <sup>+</sup> 07, KG09]                                 | No | No | No | No |
| Peste (PPR) [AER <sup>+</sup> 18,<br>M <sup>+</sup> 14, R <sup>+</sup> 18] | No | No | No | No |

Yes = Detectable by system, No = Not Detectable.

## 1.12 Gaps and open issues

There are still major gaps in the early diagnosis and monitoring of illnesses like peste in sheep, despite the increased interest in precision livestock production. In order to establish the project's relevance and necessity, this part describes the shortcomings and unresolved issues with the existing alternatives.

Only 17% of sensor fusion and AI studies deal with disease diagnosis, with the most (34.8%) concentrating on cattle and the fewest (8.7%) on sheep. This demonstrates a glaring absence of clever ways to identify serious sheep diseases like PPR. By fusing biometric sensors with artificial intelligence, the suggested smart collar seeks to close this gap and provide accurate, timely, and real-time disease detection [ALECRM<sup>+</sup>23].

Compared to pigs and cattle, sheep behavior is the subject of only 10.8% of deep learning studies. The majority make use of custom datasets, underscoring the absence of standardized data and precise descriptions of behavior. To increase accuracy and efficacy, future research should concentrate on standard datasets and cooperation with veterinarians [RRH<sup>+</sup>24].

Limitations of deep learning in veterinary diagnostics include a lack of long-term studies, selection bias, low generalizability, model interpretability problems, and a lack of diversified datasets . Practical implementation is also hampered by high prices and few resources, highlighting the need for more study and cooperation [XDW<sup>+</sup>25].

## 1.13 Conclusion

This chapter included an overview of basic livestock management practices, common sheep diseases and their symptoms, and emerging uses of artificial intelligence (AI) and the Internet of Things (IoT) in animal health monitoring. We were able to ascertain the field's present limitations and advancements by looking at existing systems. These limitations highlight the need for innovative, workable, and readily available solutions that are tailored to the unique needs of sheep herding.

The architecture of the suggested smart collar system will be presented in the upcoming chapter in order to address these issues. It will set the foundation for the practical application of an intelligent and effective disease detection tool by describing the elements, technologies, and data flow required to create it.

# System Design, System Requirements, and Architecture

## 2.1 Introduction

This chapter presents a comprehensive exploration of the smart collar system designed for the early detection of "*peste des petits ruminants (PPR)*" in sheep. It begins with an overview of the system's architecture, outlining both its functional capabilities and essential non-functional requirements that ensure reliability and efficiency. We then delve into the core hardware and software components that form the backbone of this innovative solution. The chapter also details the strategic deployment of sensors, along with the methodologies for data collection, storage, and seamless transmission. Furthermore, it covers critical steps such as data pre-processing and labeling, which prepare the information for accurate analysis. A thorough explanation of the system architecture provides the foundation for understanding how these components interact. Finally, this chapter discusses the selection and training of machine learning models, feature engineering techniques, and the performance metrics used to evaluate the system's effectiveness. Together, these elements set the stage for a robust and intelligent approach to safeguarding sheep health through cutting-edge technology.

## 2.2 Functional and Non-Functional Requirements

### 2.2.1 Functional Requirements

The following requirements describe the essential functions of the smart collar system:

- **Monitoring of Temperature:** Every three seconds, the system needs to take the sheep's body temperature. An anomaly needs to be identified and documented if the temperature rises over 40°C.
- **Blood oxygen saturation and heart rate:** The MAX30102 sensor must be used by the system to detect blood oxygen saturation (SpO<sub>2</sub>) and heart rate every three seconds. An alert should be triggered by a heart rate above 120 beats per minute or a SpO<sub>2</sub> below 90%.
- **Using a Vibration Sensor to Identify Cough:** The SW-420 sensor must be used by the system to continuously detect vibrations. When several strong vibrations are felt in a brief period of time, they should be taken as a coughing episode.
- **GPS Location:** GPS coordinates must be recorded by the system every 3 seconds. An alert must be set off if the animal leaves a designated region or stays still for an extended period of time.
- **Data Logging:** Time stamping and storing all gathered data, including temperature, heart rate, SpO<sub>2</sub>, vibrations, and GPS, is required.
- **Anomaly Detection:** To find clinical indicators of the illness, the system has to evaluate data in real time.
- **Transmission via Wireless:** Critical data must be sent from the ESP32 module to a base station via the LoRa SX1278 module.
- **Charging via Battery:** The collar needs to run on battery power and have efficient cycles for both data collecting and transmission.

### 2.2.2 Non-Functional Requirements

- **Reliability:** The system must operate stably in outdoor environments despite movement and environmental conditions.
- **Scalability:** The system must be deployable on multiple sheep simultaneously.
- **Accuracy:** Temperature measurements must have a precision of  $\pm 0.5^{\circ}\text{C}$ . Heart rate and SpO<sub>2</sub> values must be accurate within 5%.
- **Energy Consumption:** Battery life must be maximized through sleep modes and spaced data transmissions.
- **Real-Time Responsiveness:** Anomalies must be detected and processed within 3 seconds of data acquisition.
- **Maintainability:** Components must be modular and easily replaceable.
- **User Interface Compatibility:** The data must be transmittable to a web interface or mobile application.



- **Environmental Resistance:** The 3D-printed enclosure must protect the electronic components from water, dust, and impacts.

## 2.3 System Overview

### 2.3.1 General objective of the system

The goal of the smart collar technology is to enable early detection of PPR disease by continuously monitoring the health of the sheep. It incorporates several sensors, including motion, GPS, oxygen level, temperature, and heart rate, and connects them to an ESP32 microprocessor. LoRa technology is used to wirelessly transfer the gathered data to a central server, where illness detection models and data processing are implemented. If anomalous patterns are found, alerts are immediately issued, enabling farmers to take immediate action.

**The overall** operation of the proposed system is illustrated in the figure [2.1](#). It is based on the acquisition of physiological and behavioral data from sensors worn by the sheep, the analysis of this data in real time, and the generation of alerts in the event of suspected disease.

**The main steps are:**

- Data collection via onboard sensors (heart rate, temperature, GPS location).
- Data transmission to a central database via wireless communication (e.g., LoRa).
- Analysis and decision-making, based on rules or an artificial intelligence model.
- Sending an alert (e.g., SMS) to the farmer if a suspected case is detected.
- Monitoring the sheep's health status (identifying a healthy or sick sheep).

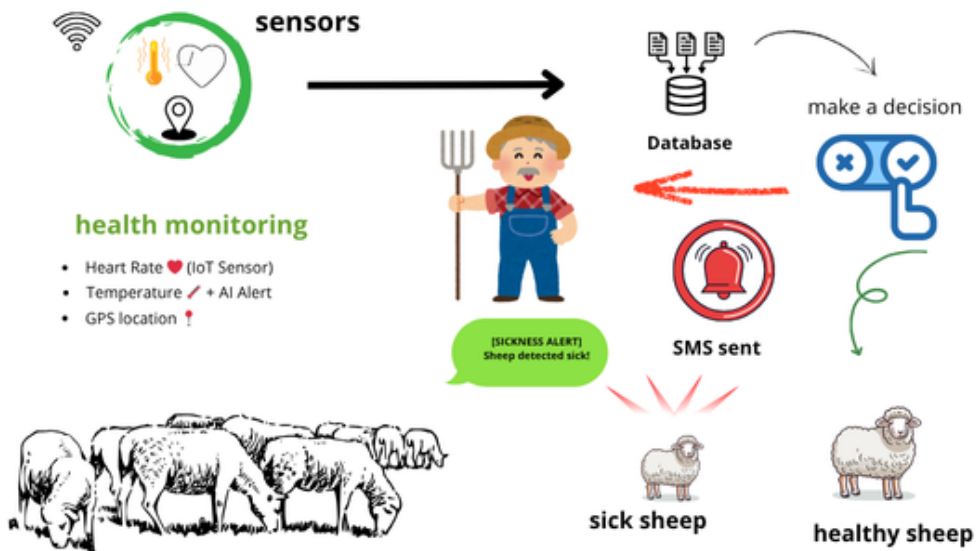


Figure 2.1 – Overview of how the smart collar works.

## 2.4 Hardware Components

The system's hardware configuration combines a number of sensors and modules to track the behavior and health of sheep in order to detect diseases like PPR early. The primary pieces of hardware are:

### 2.4.1 LoRa SX1278 with ESP32

**Function:** Serves as the main microcontroller in charge of gathering sensor data, doing initial processing, and sending data wirelessly across great distances.

**Reason for Collection:** Guarantees low-power, wide-area connectivity that is appropriate for animal monitoring in rural locations or broad fields without GSM coverage.

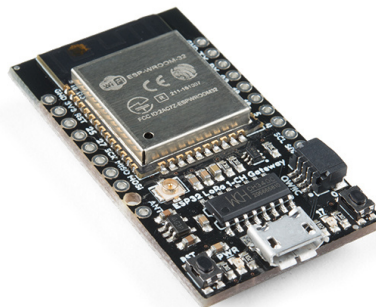


Figure 2.2 – ESP32 LoRa Microcontroller Board [W36]

### 2.4.2 MAX30102 Oxygen and Heart Rate Sensor

**Function:** Measures blood oxygen saturation ( $SpO_2$ ) and heart rate.

**Reason for Collection:** Essential for early physiological detection of PPR because respiratory distress and changed heart rates are two of its primary clinical indicators.

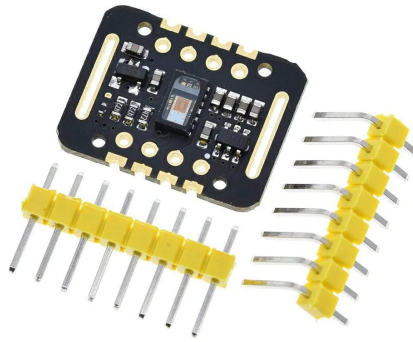


Figure 2.3 – MAX30102 Pulse Oximeter and Heart Rate Monitor [W37]

### 2.4.3 Vibration Sensor SW-420

**Function:** Identifies unusual vibrations, such as coughing.

**Reason for Collection:** Monitoring coughing aids in the early detection of respiratory disorders, such as PPR, as it is a primary symptom.



Figure 2.4 – SW-420 Vibration Sensor Module [W38]

### 2.4.4 GPS Module Neo-6M V2

**Function:** Enables real-time geolocation tracking of the sheep.

**Reason for Collection:** Any departure from typical movement patterns, like seclusion or decreased mobility, may be a sign of illness and should be closely watched.

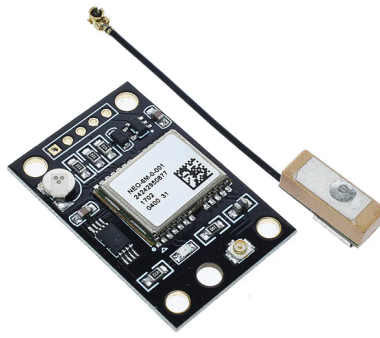


Figure 2.5 – NEO-6M GPS Module [W39]

### 2.4.5 Temperature Sensor DS18B20

**Function:** Precisely measures the sheep's body temperature.

**Reason for Collection:** Monitoring body temperature is crucial for identifying possible PPR infections because fever is a basic indicator of infection.



Figure 2.6 – Temperature Sensor DS18B20 [W40]

### 2.4.6 Rechargeable Battery

**Function:** Powers all components for continuous, portable operation.

**Reason for Collection:** Allows the wearable system to operate independently in remote and outdoor settings without requiring a continuous external power source.



Figure 2.7 – Rechargeable Battery Module [W41]

### 2.4.7 3D-printed cover for a belt

**Function:** Houses and protects the delicate electronic components.

**Reason for Collection:** Ensures the sensors and microprocessor are protected from environmental elements like dust, moisture, and physical trauma, while allowing easy placement on the sheep.

## 2.5 Software Components

### 2.5.1 Development environment

**Arduino(IDE):** A text editor for writing code, a message box, a text console, a toolbar with buttons for frequently used tasks, and a number of menus are all included in the Arduino Integrated Development Environment, also known as the Arduino Software (IDE). In order to upload programs and interact with the Arduino hardware, it connects to it.

**Google Colab:** Writing and executing Python code in your browser is possible with Google Colab, also known as Google Colaboratory, a free platform provided by Google. In particular, it enables you to execute Jupyter notebooks without worrying about the software or hardware that is present on your machine. One service that makes it easier to access computer resources and popular machine learning libraries is Google Colab.

### 2.5.2 Programming language

**C++:** an object-oriented programming (OOP) language for developing large-scale applications. With C++, programmers may create their own data types and use functions and methods to work with them. Additionally, it grants memory access and permits low-level programming, facilitating quick and effective code execution.

**Python:** Python has dynamic semantics and is a high-level, object-oriented, interpreted programming language. It is highly appealing for Rapid Application Development and for usage as a scripting or glue language to join pre-existing components because of its high-level built-in data structures, dynamic typing, and dynamic binding.

### 2.5.3 Libraries used

- **NumPy (np):** NumPy is the primary Python library for scientific computing, providing fast array operations, mathematical and statistical functions, linear algebra, Fourier transforms, random simulation, and support for masked arrays and matrices[W47].
- **Pandas (pd):** Pandas is a Python library designed for handling data sets, offering tools for analyzing, cleaning, exploring, and manipulating data [W48].
- **sklearn.preprocessing:** The sklearn.preprocessing package provides tools to transform raw feature vectors for better estimator compatibility, with standardization aiding algorithms like linear models, and robust scalers handling outliers, as seen in scaler comparisons[W49].
- **matplotlib.pyplotplt:** Seaborn serves as a state-based interface to Matplotlib, offering an implicit, MATLAB-style plotting approach. It manages figures on your screen and handles the figure GUI[W50].
- **tensorflow.keras.models:** A model that organizes layers into a unified object, incorporating features for both training and inference[W51].
- **sklearn.model selection:** Tools for model selection, such as cross validation and hyper-parameter tuning[W52].
- **sklearn.metrics:** The sklearn.metrics module provides various loss, score, and utility functions to evaluate classification performance, with some metrics needing probability estimates, confidence values, or binary decision values[W53].
- **imblearn.over sampling:** Class to perform over-sampling using SMOTE[W54].
- **tensorflow.keras.layers:** The tensorflow.keras.layers module in TensorFlow is a fundamental component for building neural networks. It provides a wide

variety of pre-built layers that can be combined to create complex models[55].

- **tensorflow.keras.callbacks:** Base class used to build new callbacks[56].
- **seaborn:** Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics[57].
- **xgboost:** XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable[58].
- **lightgbm:** LightGBM is a gradient boosting framework that uses tree based learning algorithms[59].
- **sklearn.svm:** Support vector machine algorithms[60].
- **sklearn.inspection:** module provides tools to help understand the predictions from a model and what affects them[61].
- **WiFi.h:** With the Arduino WiFi Shield, this library allows an Arduino board to connect to the internet[62].
- **HTTPClient.h:** HttpClient is a library to make it easier to interact with web servers from Arduino[63].
- **Wire.h:** This library allows you to communicate with I2C devices, a feature that is present on all Arduino boards. I2C is a very common protocol, primarily used for reading/sending data to/from external I2C components[64].
- **OneWire.h:** This library is compatible with all architectures so you should be able to use it on all the Arduino boards[65].
- **DallasTemperature.h:** Arduino library for Dallas/Maxim temperature ICs Support for DS18B20 and other Dallas/Maxim 1-Wire temperature sensors[66].
- **HeartRate.h:** The HeartRate.h Arduino library measures heart rate (BPM) from a pulse sensor (e.g., MAX30102) by detecting signal peaks and averaging beat intervals, suitable for health monitoring.
- **MAX30105.h:** The MAX30105.h Arduino library supports the MAX30105 (and MAX30102) sensor for I2C-based heart rate, SpO<sub>2</sub>, and particle measurement using red, IR, and green LEDs, with data and configuration functions.

## 2.6 System Architecture

We demonstrate the key elements of our system in this part. Using embedded sensors such as temperature, GPS, vibration (for cough detection), and the MAX30102 sensor (for heart rate and blood oxygen levels), the smart collar first gathers behavioral and physiological data from the sheep. After collection, the data is cleaned, normalized, and organized for analysis through a data preprocessing step. A variety of machine learning models are then trained using the processed data.

We test a variety of models, such as GRU, LSTM, SVM, LightGBM, and XGBoost. Every model is trained and assessed independently. Following testing, the model with the highest accuracy is chosen to be employed in the last prediction stage. The presence of "*peste des petits ruminants (PPR)*" is then predicted by this final model, which categorizes each sheep as either healthy or maybe sick, allowing for real-time alerts and preventative action.

The following diagram represents the general structure of our system:

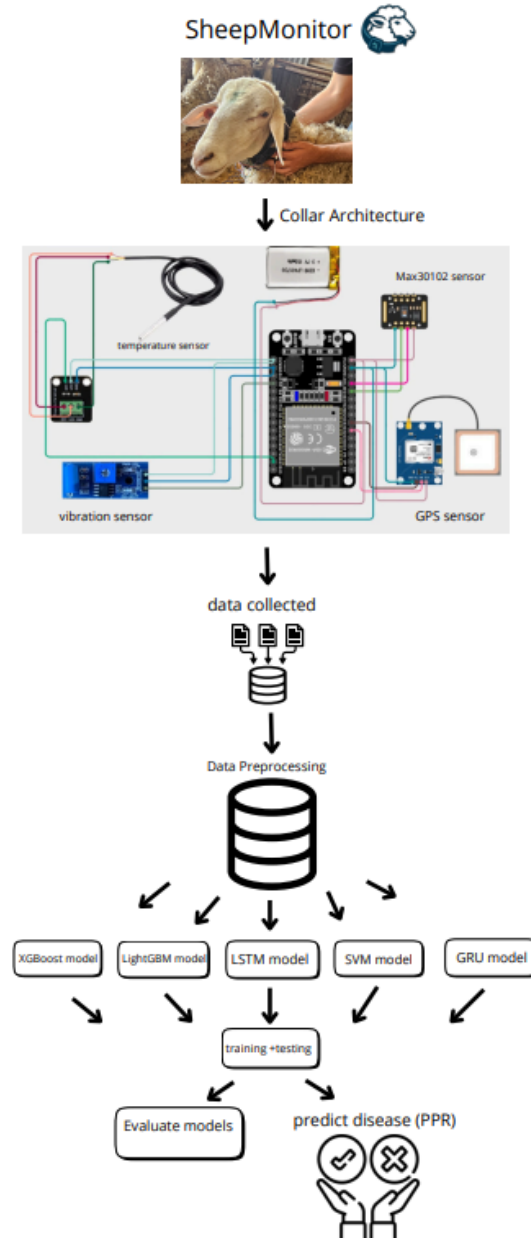


Figure 2.8 – system architecture



### 2.6.1 Deployment on the Animal:

The sheep wore a 3D-printed smart collar with all of the sensors built in. Both the animal's comfort and the electronics' safety are guaranteed by the deployment technique.

**An example of where to put a sensor:**

- The temperature sensor is positioned beneath the collar, close to the skin.
- For a stronger satellite signal, the GPS module is mounted on the collar.
- To detect coughing, a vibration sensor is positioned near the neck.
- MAX30102 gently pushed against the skin to measure heart rate and SpO<sub>2</sub>.

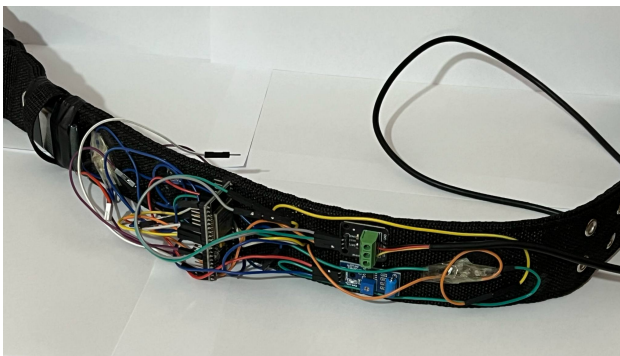


Figure 2.9 – Overview of the IoT-Based Sheep Health Monitoring Process.

### 2.6.2 Challenges and Technical Considerations:

Among the difficulties in deployment are:

- Animal movement interfering with signals.
- Waterproofing and a secure sensor attachment.
- limited battery life when exposed to the elements.

Adjusting sensor locations, insulating delicate parts, and conducting on-site testing for optimization were the solutions.

### 2.6.3 Data Collection

To monitor their health, a variety of behavioral and physiological information was collected from both healthy sheep. Sensors integrated into the smart collar recorded:

Body temperature, Heart rate, Vibration signals (for cough detection), GPS coordinates (to monitor movement and isolation).

After two hours of nonstop data collection from a single, healthy sheep, a CSV file with 1346 entries was produced. When comparing data from possibly infected sheep, this dataset acts as a baseline reference. In the future, comparable data from sheep afflicted with "*peste des petits ruminants (PPR)*" will be collected to analyze trends, detect significant variations, and improve early disease detection through machine learning analysis.

| Sensor                           | Data Collected       | Unit              | Purpose                     |
|----------------------------------|----------------------|-------------------|-----------------------------|
| Temperature                      | Body temperature     | °C                | Detect fever                |
| Heart Rate                       | Pulse rate           | bpm               | Detect stress or illness    |
| Vibration                        | Motion signal        | Binary/Analog     | Detect coughing             |
| GPS                              | Latitude & Longitude | Coordinates       | Detect isolation            |
| O <sub>2</sub> Sensor (MAX30100) | Blood oxygen level   | %SpO <sub>2</sub> | Detect respiratory problems |

Table 2.1 – List of sensors and their roles in health monitoring.

| date                | temperature | heartrate | cough |
|---------------------|-------------|-----------|-------|
| 2025-05-08 11:55:58 | 35.31       | 52        | 0     |
| 2025-05-08 11:56:01 | 35.31       | 58        | 0     |
| 2025-05-08 11:56:04 | 35.25       | 58        | 0     |
| 2025-05-08 11:56:07 | 35.31       | 63        | 0     |
| 2025-05-08 11:56:10 | 35.25       | 68        | 0     |
| 2025-05-08 11:56:14 | 35.25       | 67        | 0     |
| 2025-05-08 11:56:16 | 35.25       | 67        | 0     |
| 2025-05-08 11:56:19 | 35.25       | 67        | 0     |
| 2025-05-08 11:56:22 | 35.25       | 69        | 0     |
| 2025-05-08 11:56:25 | 35.25       | 70        | 0     |

Figure 2.10 – Field Setup During Dataset Collection for the Smart Collar System.

## 2.6.4 Data Storage and Transmission

- The ESP32 microcontroller processes and transmits the data gathered by the sensors (temperature, heart rate, vibration, and GPS).
- LoRa technology is used for wireless transmission, allowing for low-power, long-range connectivity that is ideal for isolated farming locations.
- A LoRa gateway receives the transmitted data and sends it to a central server or cloud platform.
- For analysis and visualization, the data is kept on the server in a structured format (CSV/Database).
- Users can access and track each sheep's health status in real time through a web-based dashboard.

- In the event of anomalous readings, the system is intended to notify farmers (for example, via SMS alerts), allowing for prompt action.

### 2.6.5 Data Cleaning and Preprocessing Techniques

Several pre-processing and cleaning methods were used in the smart collar data processing to guarantee the accuracy and dependability of the analyses:

#### a) Data Type Conversion:

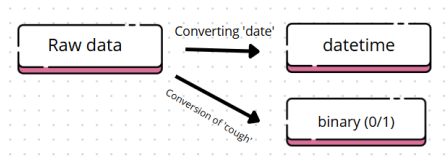


Figure 2.11 – Converting data types.

This diagram illustrates the first step of the data preprocessing process. Starting from the raw data, two types of conversions are performed. The date column is converted to a proper datetime format, making it easier to analyze and work with time-based data. The cough column, originally labeled with text values like "Yes" or "No", is converted to a binary format (1 for "Yes", 0 for "No") to facilitate machine learning and statistical analysis. These transformations ensure that the dataset is clean, consistent, and ready for further processing.

#### b) Handling Missing Values:

In this step, the dataset is checked for missing values (empty or null entries) that can affect data analysis or machine learning models. First, we inspect how many missing values exist in each column using. Then, we apply imputation techniques: Missing values in the temperature column are replaced with the mean of that column. Missing values in the heartrate column are filled with the median. Missing values in the cough column are assumed to be 0 (no cough), which is a conservative default. This step helps ensure the dataset is complete and reliable, avoiding errors or bias caused by missing data.

#### c) Data Normalization:

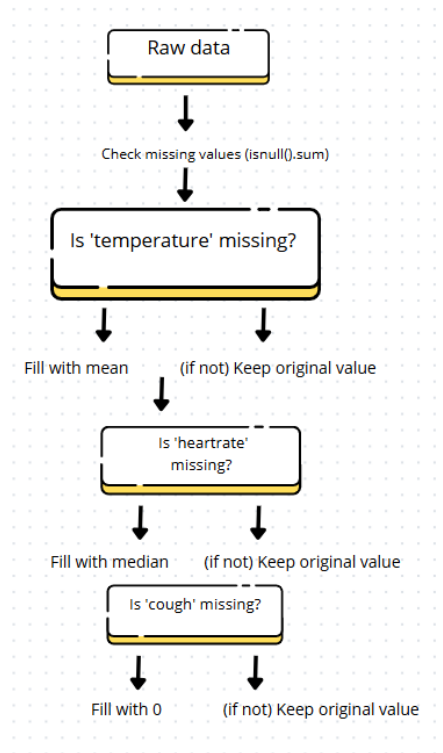


Figure 2.12 – Converting data types.

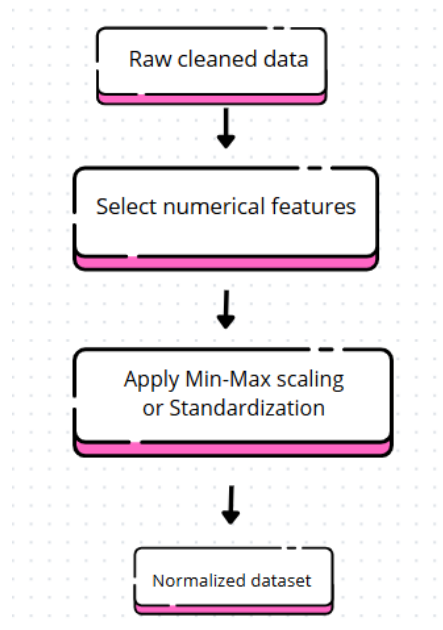


Figure 2.13 – Overview of Common Data Normalization Techniques in Preprocessing.

This step ensures that all numerical values in the dataset are brought to the same scale, which is crucial for fair comparisons and effective machine learning. Features like temperature, heartrate, or oxygen levels may have different units and ranges. Without normalization, features with larger numerical ranges could dominate the learning process, leading to biased results. We used the

following Common normalization techniques:

- **Min-Max Scaling:** Rescales data to a range between 0 and 1.

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2.1)$$

- **Z-score Standardization:** Centers the data around 0 with a standard deviation of 1. After this step, the dataset is scaled and ready for machine learning or further analysis.

$$x_{\text{std}} = \frac{x - \mu}{\sigma} \quad (2.2)$$

**d) Outlier Detection and Removal (IQR Method):**

This step is used to detect and remove extreme values that are significantly different from the rest of the data.

- We calculate the Interquartile Range (IQR) and define acceptable limits using the formula:

$$\text{IQR} = Q_3 - Q_1 \quad (2.3)$$

$$\text{Lower Bound} = Q_1 - 1.5 \times \text{IQR} \quad (2.4)$$

$$\text{Upper Bound} = Q_3 + 1.5 \times \text{IQR} \quad (2.5)$$

An outlier  $\iff x < \text{Lower Bound}$  or  $x > \text{Upper Bound}$ . Any value outside this range is considered an outlier and is removed from the dataset. This improves the data quality and helps build more accurate models.

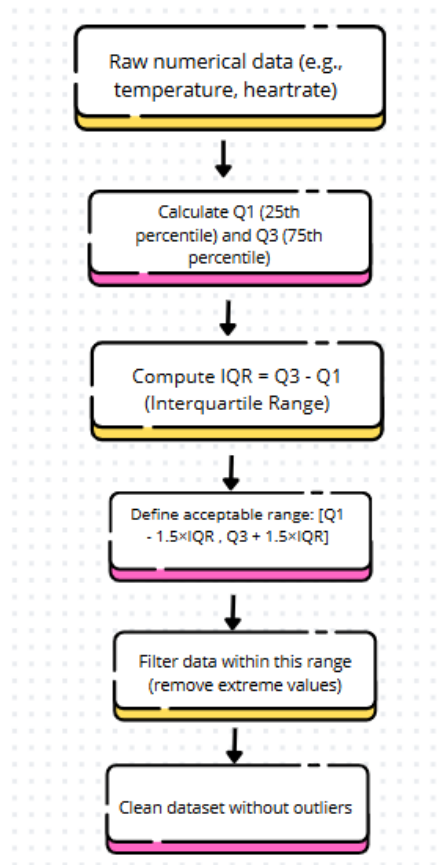


Figure 2.14 – Outlier Detection and Removal Using the IQR Method.

**e) Duplicate Removal:**

In this step, duplicate rows in the dataset are removed using the `drop_duplicates()` function. Duplicates may appear due to repeated recordings, sensor malfunctions, or data merging. By keeping only unique entries, we ensure that the dataset is clean and that no observation is overrepresented, which improves the accuracy and fairness of data analysis and machine learning models.

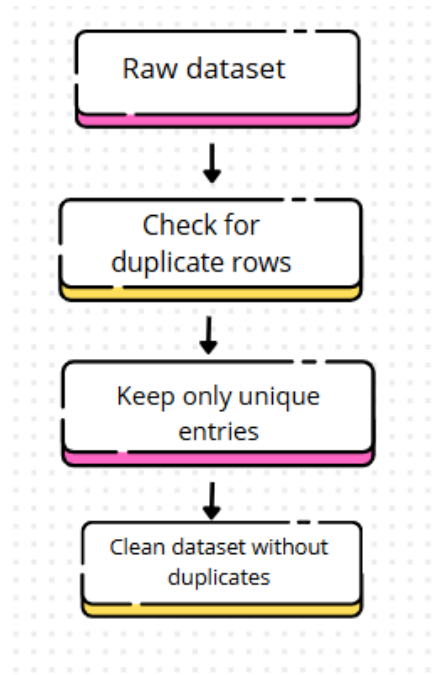


Figure 2.15 – Outlier Detection and Removal Using the IQR Method.

## 2.6.6 Labeling and Ground Truth Collection

We collaborated closely with a veterinarian to verify real disease cases in order to label the data for our PPR (Pest des Petits Ruminants) detection system in sheep. Each sheep’s behavior and symptoms were personally noted, and the sensor data was given a binary label based on diagnoses: 0: Healthy and 1: Sick (PPR confirmed). Temperature, GPS, vibration, heart rate, and oxygen saturation were among the sensors from which data was gathered and analyzed over certain time periods. This approach guarantees a trustworthy ground truth for our intelligent detection system’s training and validation.

## 2.6.7 AI Models

### a) LSTM model architecture

The input gate, forget gate, and output gate are the three gates that regulate the memory cell in LSTM systems. The information that is added to, removed from, and output from the memory cell is determined by these gates.

- The input gate regulates the data that is entered into the memory cell.
- What data is deleted from the memory cell is decided by the forget gate.
- The output gate regulates the data that the memory cell outputs.

This enables LSTM networks to learn long-term dependencies by selectively retaining or discarding information as it moves through the network. The hidden state

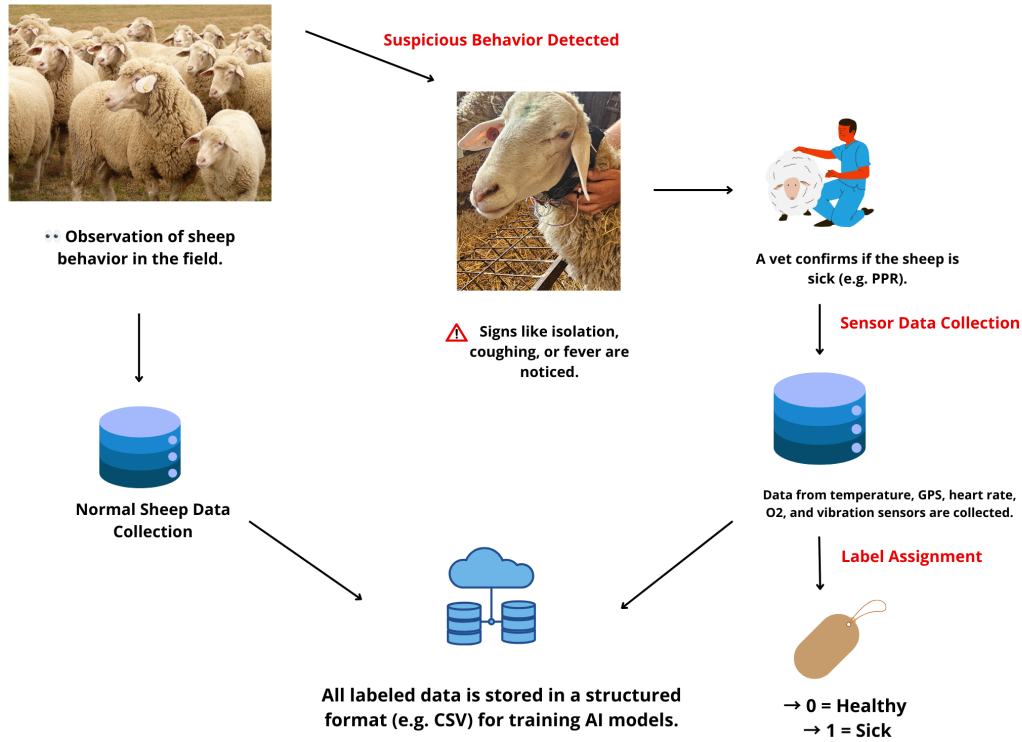


Figure 2.16 – Data Labeling for Smart Collar.

| date                | temperature | heartrate | cough | label |
|---------------------|-------------|-----------|-------|-------|
| 2025-05-08 11:55:58 | 35.31       | 52        | 0     | 0     |
| 2025-05-08 11:56:01 | 35.31       | 58        | 0     | 0     |
| 2025-05-08 11:56:04 | 35.25       | 58        | 0     | 0     |
| 2025-05-08 11:56:07 | 35.31       | 63        | 0     | 0     |
| 2025-05-08 11:56:10 | 35.25       | 68        | 0     | 0     |
| 2025-05-08 11:56:14 | 35.25       | 67        | 0     | 0     |
| 2025-05-08 11:56:16 | 35.25       | 67        | 0     | 0     |
| 2025-05-08 11:56:19 | 35.25       | 67        | 0     | 0     |
| 2025-05-08 11:56:22 | 35.25       | 69        | 0     | 0     |
| 2025-05-08 11:56:25 | 35.25       | 70        | 0     | 0     |

Figure 2.17 – Example of Labeled Dataset for Smart Collar System.

of the network functions similarly to its short-term memory. The current input, the prior concealed state, and the memory cell’s current state are used to update this memory.

LSTM architecture has a chain structure that contains four neural networks and



different memory blocks called cells[W42].

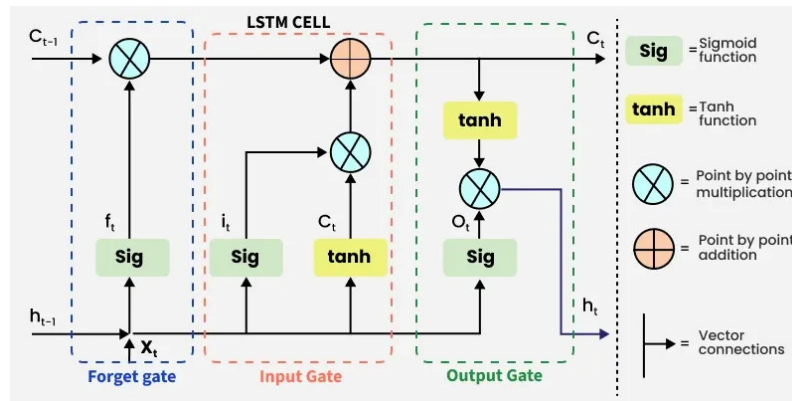


Figure 2.18 – The architecture of the LSTM method [W42]

### b) XGBOOST model architecture

The scalable and effective implementation of gradient boosted decision trees in the XGBoost architecture makes it stand out. It has functions including missing data management, regularization to avoid over-fitting, and an adaptable methodology that lets users design their own optimization goals and standards. The accuracy and resilience of the predictive model are enhanced by these characteristics[43].

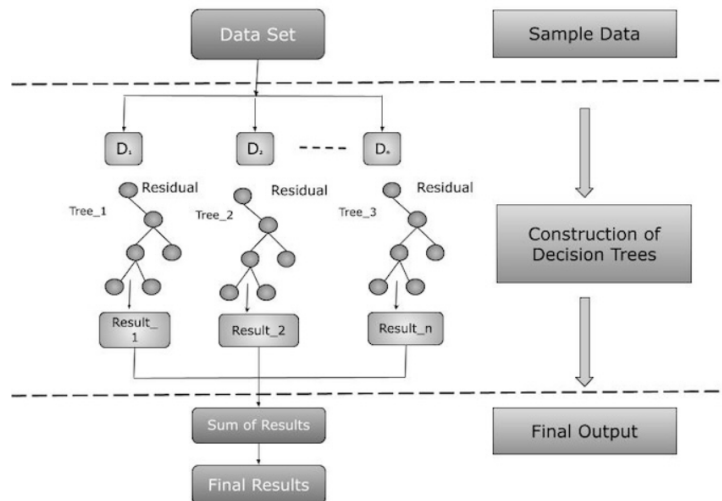


Figure 2.19 – The architecture of the XGBOOST method [43]

### c) SVM model architecture

A supervised machine learning approach for classification and regression applications is called Support Vector Machine (SVM). It maximizes the margin between data points of distinct classes by identifying the ideal hyperplane to divide them.

**Feature Vectors:** Each input sample is represented as a vector of features:

$$\mathbf{x} = [x_1, x_2, \dots, x_n] \quad (2.6)$$

**Support Vectors:** These are the critical data points closest to the decision boundary. They determine the optimal hyperplane.

**Kernel Function:** Used to map input data into a higher-dimensional space to handle non-linearly separable data. Common kernels:

- Linear:  $K(x, x_i) = x^\top x_i$
- Polynomial:  $K(x, x_i) = (x^\top x_i + c)^d$
- Radial Basis Function (RBF):  $K(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$
- Sigmoid:  $K(x, x_i) = \tanh(\alpha x^\top x_i + c)$

**Optimization (Training):** The algorithm solves a convex optimization problem to find the Lagrange multipliers  $\alpha_i$  and bias term  $b$ , maximizing the margin between classes.

**Decision Function:** Once trained, the model predicts using the decision function:

$$f(x) = \sum_{i=1}^n \alpha_i y_i K(x, x_i) + b \quad (2.7)$$

The sign of  $f(x)$  determines the predicted class.

**Predicted Class:** The model outputs either class +1 or -1 (or multiclass via extension), depending on the sign of the decision function.

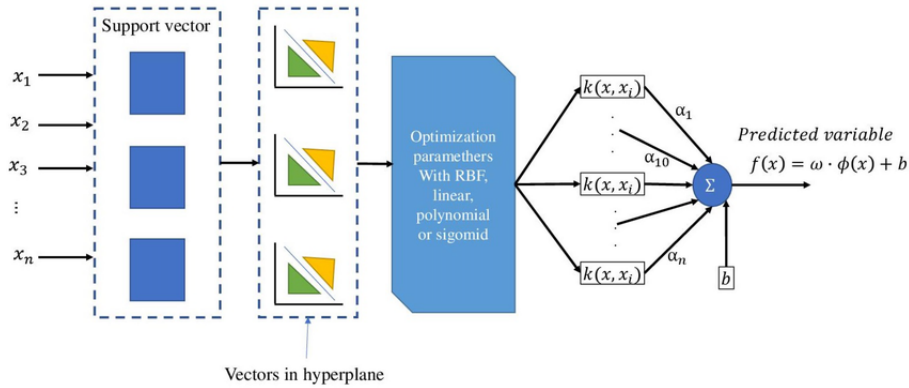


Figure 2.20 – The architecture of the SVM method [44]

**d) LightGBM model architecture**

LightGBM (Light Gradient Boosting Machine) is a high-performance, distributed, gradient boosting framework based on decision tree algorithms. It is designed for speed and efficiency, particularly on large datasets and high-dimensional data.

**Input Data:** The model receives tabular data with features  $X = [x_1, x_2, \dots, x_n]$  and corresponding labels  $y$ .

**Histogram-based Decision Tree:** Continuous features are discretized into histogram bins. This improves memory efficiency and reduces computation time.

**Gradient Boosting Framework:** Trees are added sequentially. Each tree is trained to correct the errors (gradients) of the previous ensemble.

**Leaf-wise Tree Growth Strategy:** LightGBM grows trees leaf-wise, always splitting the leaf with the largest loss reduction. This often leads to better accuracy compared to level-wise growth.

**Loss Optimization:** The model uses gradient (first-order derivative) and Hessian (second-order derivative) statistics to optimize the loss function.

**Output Prediction:** The final prediction is the sum of the outputs from all the trees:

$$\hat{y} = \sum_{t=1}^T f_t(x) \quad (2.8)$$

where  $f_t$  is the prediction function of the  $t$ -th tree.

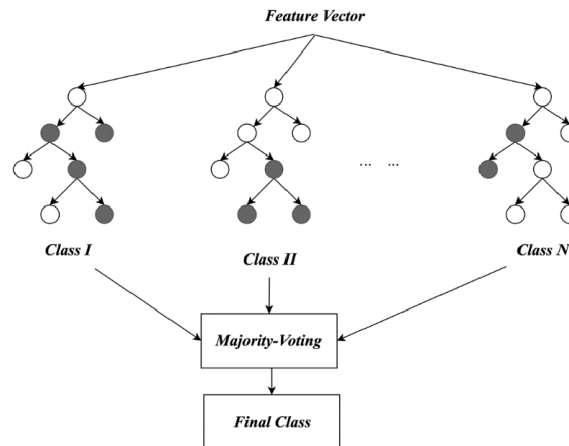


Figure 2.21 – The architecture of the LightGBM method [45]

**e) GRU model architecture**

The Gated Recurrent Unit (GRU) is a type of recurrent neural network (RNN) that uses gating mechanisms to control the flow of information. It aims to capture long-term dependencies in sequences while being computationally efficient. At each time step  $t$ , the GRU performs the following operations:

• **Input:**

$$x_t \text{ (input vector), } h_{t-1} \text{ (previous hidden state)} \quad (2.9)$$

• **Update Gate:**

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (2.10)$$

• **Reset Gate:**

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (2.11)$$

• **Candidate Hidden State:**

$$\tilde{h}_t = \tanh(W_h x_t + U_h(r_t \odot h_{t-1})) \quad (2.12)$$

• **Final Hidden State:**

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (2.13)$$

Where:

- $\sigma$ : Sigmoid activation function
- $\odot$ : Element-wise multiplication
- $W_*, U_*$ : Trainable weight matrices

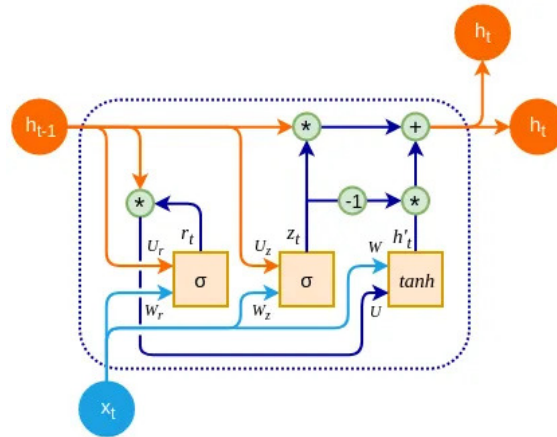


Figure 2.22 – The architecture of the GRU method [46]

## 2.7 Model Training and Validation

To properly evaluate the model and prevent overfitting, we adopted a two-level data splitting approach. First, we split the entire dataset into 80% for training and validation, and 20% for testing. Then, within the 80% training portion, we applied a further validation split of 0.2, meaning that 16% of the total data was used for validation and 64% for actual training.

This strategy ensures that model performance is monitored on unseen validation data during training, while the final test set remains completely separate for unbiased performance evaluation after training is complete.

## 2.8 Symptom-to-Sensor Mapping

Converting biological symptoms into observable and quantifiable parameters is crucial for the smart collar system's ability to identify clinical indicators of "*peste des petits ruminants (PPR)*". The primary clinical signs of PPR and the kinds of sensors that can be used to identify them are mapped in table 2.2. Additionally, it outlines the ideal location for the sensor to guarantee precision and animal comfort.

| Clinical Symptom            | Observable or Measurable Indicator                | Recommended Sensor                   | Placement / Integration                |
|-----------------------------|---------------------------------------------------|--------------------------------------|----------------------------------------|
| High fever                  | Body temperature > 40°C                           | Temperature sensor (DS18B20)         | neck contact                           |
| Coughing                    | Vibrations, neck jerks, repetitive sounds         | Vibration sensor (SW-420)            | On the collar, throat side             |
| Rapid / Difficult breathing | Changed breathing rate, low oxygenation           | SPO2 + Heart rate sensor (MAX30102)  | Collar (near jugular artery or ear)    |
| Nasal discharge             | Persistent localized moisture                     | Humidity sensor (capacitive)         | Extension on snout / harness           |
| Tearing                     | Moisture around eyes                              | Miniature humidity sensor (optional) | Less integrated – optional extension   |
| Diarrhea                    | Rapid weight loss / disturbed intestinal activity | Weight sensor (scale under trough)   | Feeding trough – not collar-integrated |
| Loss of appetite            | Fewer visits to trough, weight loss               | Scale + RFID (feeding activity)      | External infrastructure                |
| Mouth ulcers / lesions      | Pain → excessive scratching, agitation            | Accelerometer (MPU6050)              | On the collar                          |
| Weakness / Prostration      | Low activity, few movements                       | Accelerometer + GPS                  | On the collar                          |
| Herd isolation              | Abnormal movement, away from group                | GPS (Neo-6M)                         | On the collar                          |

Table 2.2 – Transformation of Clinical Symptoms into Measurable Data.

Non-invasive sensors built into a smart collar can efficiently convert a number of PPR clinical symptoms into quantifiable data, but some signs are still challenging to record without endangering the animal's wellbeing or necessitating further infrastructure. It can be difficult to continuously monitor symptoms like mouth ulcers or nasal discharge, for instance, because they may not show obvious external behaviors or may call for sensors to be positioned in delicate locations like the mouth or snout. Similarly, in order to identify changes in weight or feeding habits, diarrhea frequently necessitates the use of external equipment, such as feeding station balances.

## **2.9 Conclusion**

The functional and non-functional needs of the system, as well as its general architecture, were thoroughly covered in this chapter. We described the chosen hardware and software elements intended to facilitate consistent and dependable data collection using the right sensors. From sensor deployment to data transmission, storage, and cleaning, the system architecture covers all of the important phases of data processing. Finding pertinent features through annotated data gathering helped direct the model selection procedure. Lastly, strict performance criteria were used to train and validate the model, guaranteeing the correctness and resilience of the system. This scientific approach provides a strong basis for creating a clever and practical solution for monitoring the health of sheep.

## Implementation, Evaluation and Results

### 3.1 Introduction

The experimental validation of the suggested sheep health monitoring system is the main topic of this chapter. Its primary goal is to evaluate the system's functionality, precision, and dependability in identifying early illness indicators in practical settings.

We start by outlining the experimental configuration and the many test scenarios intended to assess the system in both field and controlled settings. The system's technical performance is then examined in terms of stability, data transfer, and responsiveness. Next, using measures like precision, recall, and detection rate, we assess the AI model's performance in terms of detection accuracy and categorization. The functioning and adaptability of the system are verified in real-world farming scenarios through case studies and field testing. Key findings and lessons from the experiments are highlighted in the results interpretation. We also go over the primary implementation issues, such as hardware limitations and environmental fluctuations.

With an emphasis on data privacy, animal welfare, and the future of smart farming, the chapter concludes by discussing the ethical and societal ramifications of adopting AI technologies in livestock health monitoring. When taken as a whole, these assessments offer a thorough understanding of the system's preparedness for implementation and future development.

## 3.2 Experimental Setup and Scenarios

### 3.2.1 Experimental Setup

The smart collar system for identifying "*peste des petits ruminants (PPR)*" in sheep was implemented in phases, beginning with virtual simulation and concluding with field testing.

#### a) Simulation on Wokwi

The first step consisted of conducting a virtual simulation using the Wokwi platform. This simulation aimed to validate the basic operation of the sensors used in the system, including:

- Temperature sensor (DS18B20)
- GPS module (for location and movement tracking)
- Vibration sensor SW-420 (for cough detection)
- MAX30102 sensor (for heart rate and oxygen level monitoring).

The simulation helped test the logic, data acquisition, and initial communication functionalities before moving to the physical implementation.

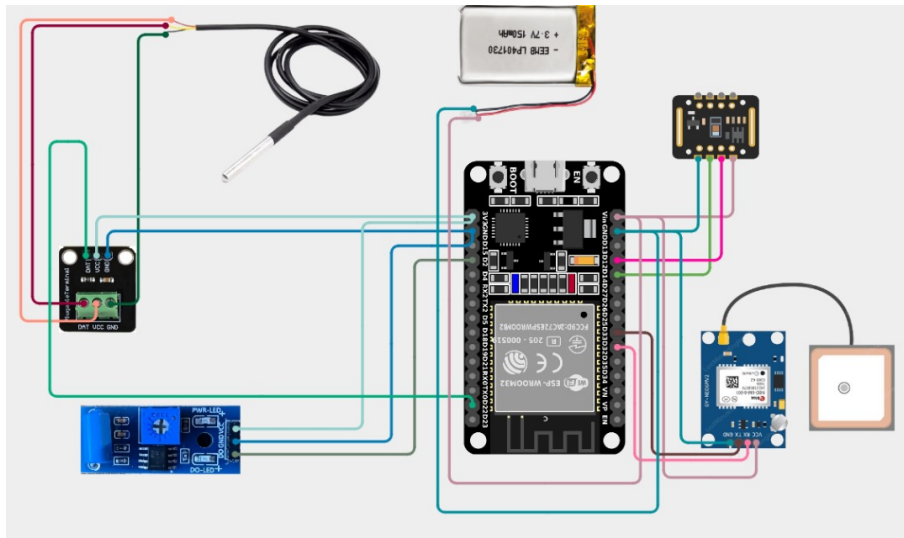


Figure 3.1 – ESP32-Based Health Monitoring System Circuit Simulation



### **b) Physical Implementation of the Smart Collar**

After validating the concept through simulation, the system was physically assembled into a wearable smart collar that integrates all the sensors listed above. The collar was designed to be comfortably worn around the neck of the sheep without interfering with its normal behavior.

The firmware was developed using the Arduino IDE, and included the following functionalities:

- Sensor data acquisition and management
- Wi-Fi connectivity for sending data to a remote platform
- LoRa module for long-range communication in rural environments
- Data formatting and automatic transmission

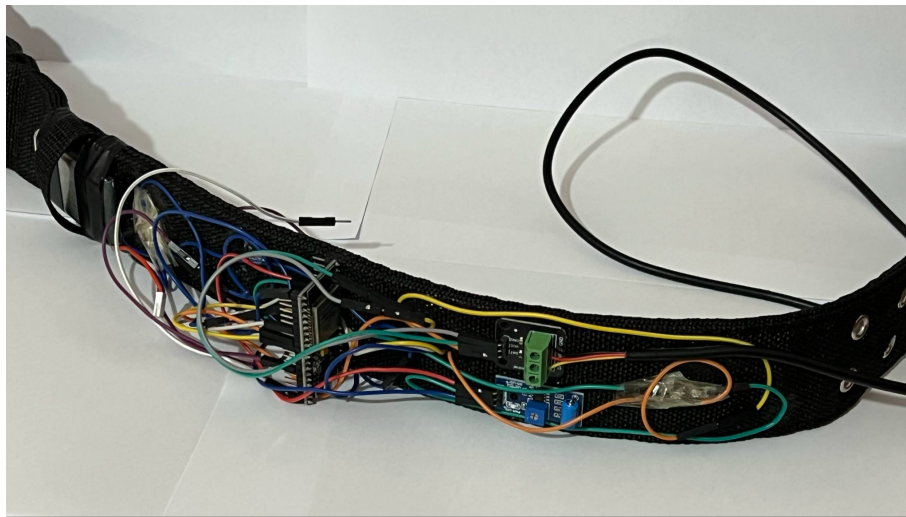


Figure 3.2 – Smart Collar Prototype for Early Detection of PPR in Small Ruminants

### **c) Data Transmission**

The data collected by the sensors is transmitted via Wi-Fi or LoRa, depending on the connectivity conditions, to a remote platform for further processing and analysis. This process is carried out automatically and periodically, enabling continuous health monitoring of the sheep.

### **d) Field Testing**

To ensure proper operation of the system, a real-world test was conducted at a farm called Fermicolle. The smart collar was tested on a healthy sheep, with the objectives of verifying:

- The physical stability and comfort of the collar on the animal

- The accuracy of the collected sensor data
- The reliability of data transmission under real conditions

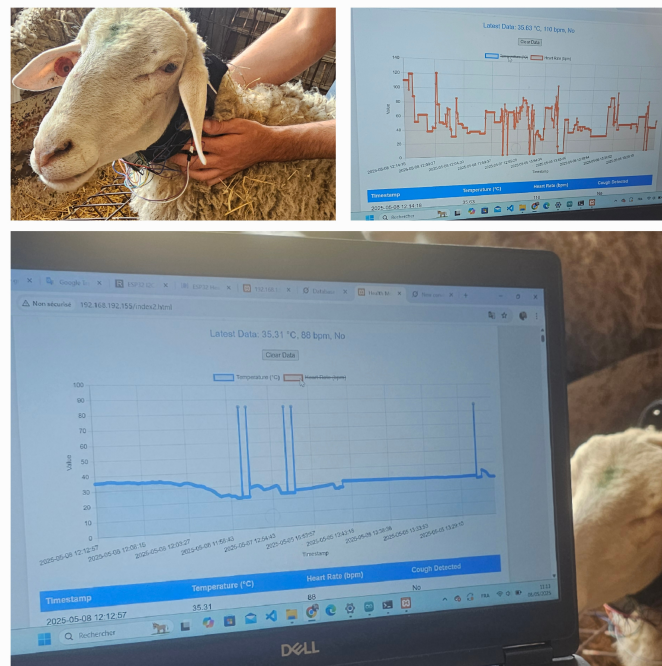


Figure 3.3 – Field Test of the Smart Collar on a Sheep

The results of this initial field test confirmed the overall functionality and reliability of the system in a real farm environment.

### 3.2.2 Experimental Scenarios

#### Scenario 1: Behavior of a Healthy Sheep

| Sensor                               | Measured Parameter               | Expected Normal Range            | Remarks                                    |
|--------------------------------------|----------------------------------|----------------------------------|--------------------------------------------|
| Temperature Sensor (DHT22 / DS18B20) | Body Temperature                 | 38.5°C – 40.0°C                  | Normal temperature range for healthy sheep |
| GPS Module                           | Position / Movement              | Regular movement within the herd | Sheep moves normally and is not isolated   |
| Vibration Sensor (SW-420)            | Vibration (Cough Detection)      | No unusual vibrations            | No signs of coughing or agitation          |
| MAX30102                             | Heart Rate                       | 70 – 90 bpm                      | Normal heart rate for a resting sheep      |
| MAX30102                             | Blood Oxygen (SpO <sub>2</sub> ) | 95% – 100%                       | Healthy blood oxygen saturation            |

Table 3.1 – Healthy Sheep Normal Sensor Values.

#### Scenario 2: Infected Sheep – Abnormal Sensor Values (PPR case)

| Sensor                               | Measured Parameter          | Abnormal Values (PPR)              | Remarks                                         |
|--------------------------------------|-----------------------------|------------------------------------|-------------------------------------------------|
| Temperature Sensor (DHT22 / DS18B20) | Body Temperature            | > 40.5°C                           | High fever is a common symptom of PPR           |
| GPS Module                           | Position / Movement         | Minimal or no movement             | Infected sheep isolate themselves and move less |
| Vibration Sensor (SW-420)            | Vibration (Cough Detection) | Frequent vibrations (cough spikes) | Dry or wet coughing is often observed           |

|          |                                  |           |                                                 |
|----------|----------------------------------|-----------|-------------------------------------------------|
| MAX30102 | Heart Rate                       | > 100 bpm | Elevated heart rate indicates stress or illness |
| MAX30102 | Blood Oxygen (SpO <sub>2</sub> ) | < 90%     | Hypoxia due to respiratory issues               |

Table 3.2 – Infected Sheep Abnormal Sensor Values (PPR case)

### Scenario 3: Sheep in Alert or Recovery Phase.

| Sensor                               | Measured Parameter               | Observed Values (Borderline)  | Remarks                                                    |
|--------------------------------------|----------------------------------|-------------------------------|------------------------------------------------------------|
| Temperature Sensor (DHT22 / DS18B20) | Body Temperature                 | 40.0°C – 40.5°C               | Mild fever, possible early-stage infection or recovery     |
| GPS Module                           | Position / Movement              | Moderate decrease in movement | Sheep remains with the flock but shows less activity       |
| Vibration Sensor (SW-420)            | Vibration (Cough Detection)      | Occasional spikes             | Infrequent coughing or throat irritation                   |
| MAX30102                             | Heart Rate                       | 90 – 100 bpm                  | Slightly elevated, possibly due to stress or mild symptoms |
| MAX30102                             | Blood Oxygen (SpO <sub>2</sub> ) | 90% – 94%                     | Slight drop in oxygen saturation, early or fading distress |

Table 3.3 – Sheep in Alert.

## 3.3 Models hyperparameters

Whichever kind of artificial intelligence application, achieving acceptable performance depends critically on the model parameterization stage. We conducted multiple experiments and experimentally tuned the parameters to get the best possible results. We carefully examined our data and came up with the following numbers to optimize our systems:

### 3.3.1 LSTM Model Parameters

- **Input shape:** (Timestamps, Features) (1, 12) for 1 time steps and 12 features.
- **Number of LSTM layers:** 1
- **Units per LSTM layer:** 50
- **Dropout rate:** 0.2
- **Activation function:** ReLU (in dense layers), tanh (in LSTM cells)
- **Output layer:** Dense with sigmoid activation for binary classification (0 vs 1)
- **Loss function:** Binary Crossentropy
- **Optimizer:** Adam
- **Learning rate:** 0.001
- **Batch size:** 32
- **Number of epochs:** 50
- **Validation split:** 20%

### 3.3.2 SVM Model hyperparameters

- **Kernel function:** Radial Basis Function (RBF)
- **Regularization parameter (C):** 1.0
- **Kernel coefficient (gamma):** 'scale'
- **Decision function shape:** One-vs-Rest (OvR)
- **Tolerance (tol):** 0.001
- **Maximum iterations:** -1
- **Feature scaling:** StandardScaler

### 3.3.3 XGBOOST Model hyperparameters

- **max\_depth:** 5
- **learning\_rate:** 0.1
- **n\_estimators:** 100
- **random state:** 42
- **objective:** binary:logistic
- **eval\_metric:** logloss

### 3.3.4 LightGBM Model hyperparameters

- **boosting\_type:** gbdt (Gradient Boosting Decision Tree)

- **objective:** binary (pour classification binaire)
- **learning\_rate:** 0.1
- **num\_leaves:** 31
- **max\_depth:** -1
- **n\_estimators:** 100
- **subsample:** 1
- **colsample\_bytree:** 1
- **min\_child\_samples:** 20
- **min\_child\_weight:** 0.001
- **lambda\_l1:** 0

### 3.3.5 GRU Model hyperparameters

- **Input Shape:** (1,12)
- **Hidden size:** 50 units
- **Number of layers:** 5
- **Dropout:** 0.2
- **Bidirectional:** not used
- **Batch size:** 32
- **Learning rate:** 0.001
- **Optimizer:** Adam
- **Loss function:** Binary Cross Entropy (ou autre selon la tâche)
- **Epochs:** 50 (nombre total de passages sur l'ensemble d'entraînement)

## 3.4 Dataset

Before discovering a dataset with sufficient pertinent and organized clinical data about the "*peste des petits ruminants (PPR)*" illness in sheep and goats, we conducted a thorough search. 161 samples with a range of clinical symptoms and RT-qPCR test results were included in the initial dataset, which was gathered from six districts in Tanzania's northern region. A data synthesis approach (CTGAN) was used to create a balanced and statistically valid dataset of 21,167 samples, but only 12 positive examples were found because of class imbalance. We were able to develop and assess machine learning models for the early identification and prediction of PPR disease in small ruminants thanks to this enriched dataset.

### 3.5 Evaluation metrics

- **Accuracy :** In classification, the accuracy metric is an important measurement to validate the model. Accuracy, or the rate of correct predictions, corresponds to the proportion of correct predictions among all predictions made. It is the number of true positives and true negatives predicted by the model divided by the total number of data points. The accuracy is calculated using the following formula:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.1)$$

Where:

- $TP$  = True Positives
- $TN$  = True Negatives
- $FP$  = False Positives
- $FN$  = False Negatives

This formula helps in determining how well the model is performing by considering both the correctly identified positive and negative instances.

- **Precision :** Precision is a measure of user performance in classification and automatic learning that evaluates the model's ability to correctly identify positive examples among all positive predictions. The names of the words and the faux positives are divided by the names of the voices to calculate the precision. In these terms, it is necessary to provide exact positive results for all positive predictions. The level of accuracy indicates that the model is capable of producing a viable combination of faux positives, which is generally desirable. However, high precision can increase the false negatives because this model has a negative class that is positive. The precision is calculated using the following formula:

$$\text{Precision} = \frac{TP}{TP + FN} \quad (3.2)$$

- **Recall :** In classification and machine learning, Recall is a performance metric used to evaluate the ability of a model to identify all positive examples among all positive real examples.

The calculation is obtained by dividing the number of true positives by the sum of true and false negatives. In other words, it calculates the appropriate positive rate for each positive real-world example. The recall is calculated using the following formula:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.3)$$

- **F1-score :** The F1 score is a harmonic average of precision and recall. It is used to consider both precision and recall, placing more emphasis on uneven classes. A value of 1 on the F1-score indicates perfect classification, while a value of 0 indicates poor classification. Therefore, a higher F1 score means that the classification model performs better in terms of precision and recall. The F1-score is calculated using the following formula:

$$\text{Recall} = \frac{2.(Precision.Recall)}{Precision + Recall} \quad (3.4)$$

### 3.6 Quantitative Findings

In this section, we will discuss the outcomes of the five developed models: the LSTM model, the SVM model, the XGBoost model, the GRU model, and the LightGBM model.

**LSTM model:** We examined the confusion matrix derived from the test set in order to assess how well the LSTM model classified the health state of sheep (healthy vs. sick). By contrasting the actual and anticipated labels, this matrix offers comprehensive insights into the classification findings. It enables us to determine the number of false positives and false negatives, which are critical in a health monitoring setting, as well as how well the model differentiates between positive (ill) and negative (healthy) situations. The confusion matrix's findings are shown below.

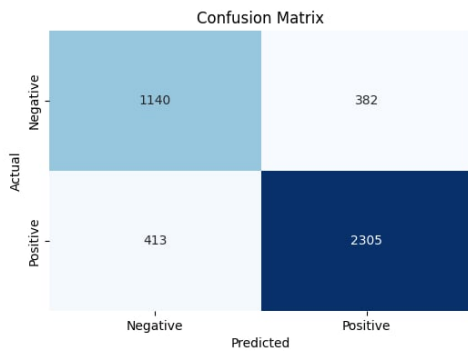


Figure 3.4 – The confusion matrix for LSTM model.

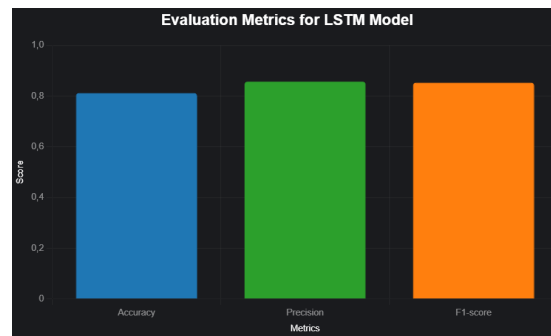


Figure 3.5 – The evaluation metrics plots for LSTM model.

**SVM model:** We employed a confusion matrix based on the predictions made on the test set to evaluate how well the Support Vector Machine (SVM) model classified the health status of sheep. This matrix demonstrates how well the model



can distinguish between animals that are ill and those that are well. It offers a thorough analysis of false positives, false negatives, real positives, and true negatives. Understanding the SVM model's diagnostic reliability is crucial, particularly in a real-world cattle health monitoring scenario. Below is the confusion matrix.

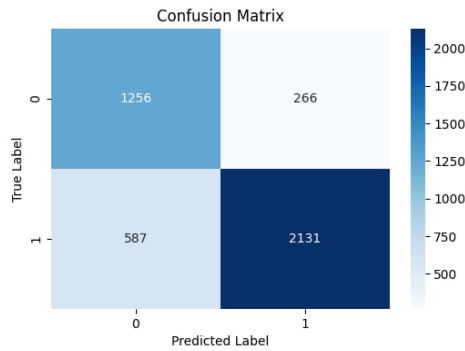


Figure 3.6 – The confusion matrix for SVM model.

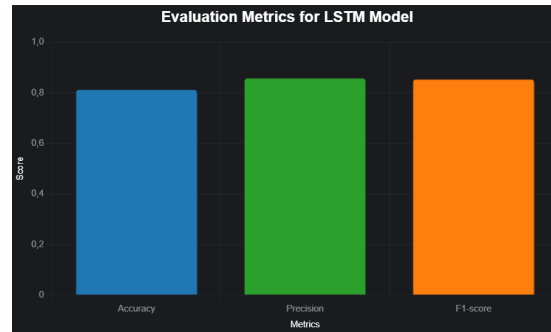


Figure 3.7 – The evaluation metrics plots for SVM model.

**XGBoost model:** A confusion matrix based on predictions made on the test dataset was created in order to assess how well the XGBoost classifier performed in identifying sick and healthy sheep. The model's ability to differentiate between the two classes is clearly shown in this matrix. It displays the amount of misclassifications (false positives and false negatives) as well as the number of accurate forecasts (true positives and true negatives). A useful technique for evaluating XGBoost's robustness and dependability in the context of cattle disease detection is the confusion matrix. The figure below displays the findings.

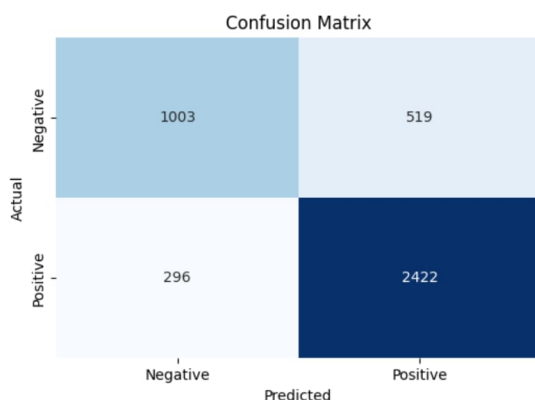


Figure 3.8 – The confusion matrix for XGBoost model.

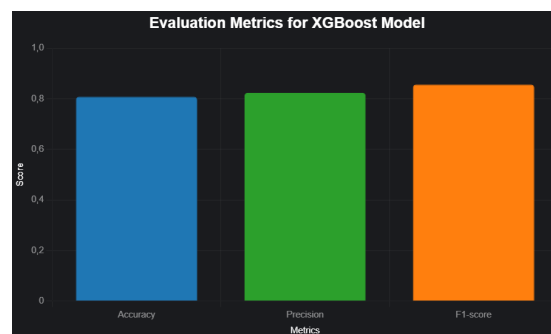


Figure 3.9 – The evaluation metrics plots for XGBoost model.

**GRU model:** Using a confusion matrix derived from the test dataset's predictions, we assessed the GRU (Gated Recurrent Unit) model's classification performance.

We can see how successfully the model distinguishes between healthy and sick sheep thanks to this matrix. By contrasting expected labels with real ground truth values, it indicates the proportion of accurate and inaccurate classifications. To evaluate the GRU model's performance in an actual animal health monitoring situation, this analysis is crucial. Below is the confusion matrix.

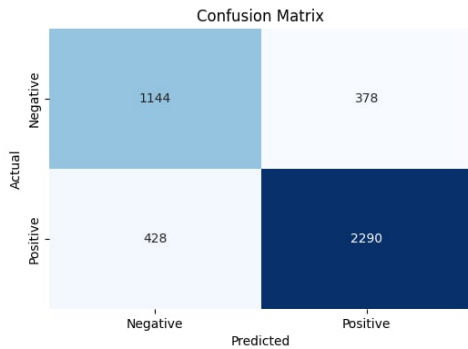


Figure 3.10 – The confusion matrix for GRU model.

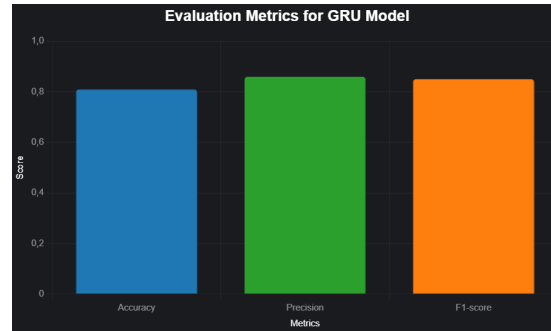


Figure 3.11 – The evaluation metrics plots for GRU model.

**LightGBM model:** We examined the LightGBM model's confusion matrix produced from the test data in order to evaluate its classification performance. We can assess the model's precision in identifying sick and healthy sheep thanks to this matrix, which offers a thorough comparison of actual and anticipated class labels. We may have a better understanding of the LightGBM model's advantages and disadvantages in illness detection tasks by looking at the true positives, true negatives, false positives, and false negatives. The following is the confusion matrix.

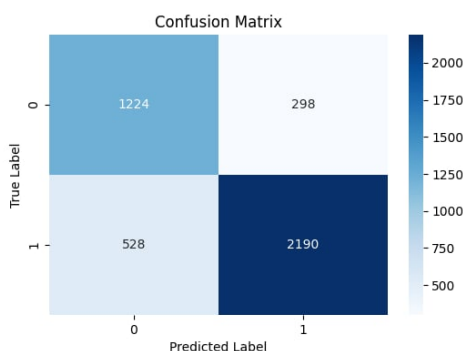


Figure 3.12 – The confusion matrix for LightGBM model.

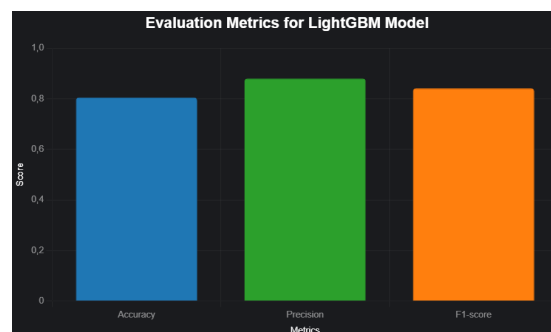


Figure 3.13 – The evaluation metrics plots for LightGBM model.

### 3.7 Comparative Analysis

The performance of five machine learning models—LSTM, GRU, SVM, XGBoost, and LightGBM—is compared in this section. The evaluation metrics acquired for every model on the test dataset are compiled in the following table:

| Metric / Model           | LSTM | GRU  | SVM  | XGBoost | LightGBM |
|--------------------------|------|------|------|---------|----------|
| <b>Accuracy</b>          | 0.81 | 0.81 | 0.80 | 0.81    | 0.80     |
| <b>Testing Precision</b> | 0.86 | 0.85 | 0.89 | 0.82    | 0.89     |
| <b>Testing Recall</b>    | 0.84 | 0.84 | 0.78 | 0.90    | 0.80     |
| <b>Testing F1 Score</b>  | 0.85 | 0.85 | 0.83 | 0.85    | 0.84     |

Table 3.4 – Comparison of model performance.

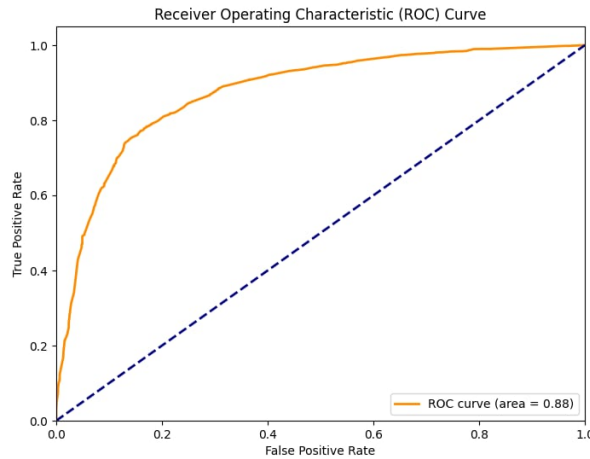


Figure 3.14 – The LSTM Model’s ROC Curve and AUC Score.

Regarding accuracy, precision, and F1-score, the LSTM model outperforms the other four models (GRU, SVM, XGBoost, and Transformer).

Specifically, the LSTM model demonstrates the highest testing accuracy, indicating its strong generalization capability. Furthermore, its precision and F1-score suggest that it is not only accurate in its predictions but also consistent in correctly identifying both positive and negative classes. This superior performance may be attributed to the LSTM’s ability to effectively capture temporal dependencies in the sequence data, which is critical in our classification task. In contrast, the other models, especially SVM and XGBoost, show lower performance, likely due to their limited handling of sequential patterns compared to LSTM.

### 3.8 System Development and Deployment Challenges

The development of the smart collar system for early detection of Peste des Petits Ruminants (PPR) in sheep involved several technical and practical challenges, from hardware integration to field deployment:

- (a) **Sensor Integration and Compatibility:** Combining various sensors (temperature, GPS, vibration, MAX30102) on a single ESP32 board was complex due to differences in communication protocols and voltage requirements. It required precise wiring, resistor calibration, and software optimization.
- (b) **Power Management:** The use of multiple sensors and dual communication modules (Wi-Fi and LoRa) led to high power consumption. Energy-saving strategies such as sleep modes and transmission frequency control were implemented to extend battery life.
- (c) **Physical Design and Environmental Resistance:** Designing a durable, lightweight, and weather-resistant collar was challenging. The 3D-printed case needed refinement to ensure comfort for the sheep and accurate sensor readings.
- (d) **Real-World Testing:** Field tests on healthy sheep (Fermicolle farm) revealed issues with data stability due to animal movement and environmental noise, which required additional signal filtering techniques.
- (e) **Lack of Existing Sensor-Based Datasets:** There is a shortage of publicly available datasets containing physiological sensor data (e.g., temperature, heart rate) for sheep. This limited the ability to train machine learning models and necessitated manual data collection through simulated and real scenarios.

### 3.9 Ethical and Societal Implications

The deployment of a smart collar system for early detection of Peste des Petits Ruminants (PPR) raises important ethical and social considerations:

- (a) **Animal Welfare and Comfort:** The device was designed with soft, lightweight materials to avoid harming or stressing the sheep. Sensors operate continuously and non-invasively, preserving the animals' natural behavior. All tests were supervised to ensure no negative effects on animal health.
- (b) **Data Privacy and Security:** Although the system collects animal data, farmers' privacy is a priority. Secure data transmission (e.g., with encryption) and informed consent are required before sharing data with veterinarians or research institutions.

- (c) **Socioeconomic Benefits for Farmers:** Early disease detection reduces treatment costs and improves herd productivity. Real-time monitoring can also save labor and time, especially in large or remote farms.
- (d) **Responsible Use of Technology:** The system is intended to support—not replace—veterinary care. Ethical use involves maintaining regular health protocols (vaccination, checkups) while using the collar as a complementary decision-support tool.

### 3.10 Conclusion

This chapter has provided a comprehensive evaluation of the smart collar system designed for early detection of PPR in sheep, demonstrating its functionality, reliability, and potential for real-world application. Through rigorous experimental setups, including virtual simulations on Wokwi and field testing at Fermicolle farm, the system proved its capability to monitor key health indicators such as temperature, heart rate, and movement with high accuracy and stability. The AI models, particularly the LSTM model, achieved superior performance with an accuracy of 0.81, a precision of 0.85, and an F1-score of 0.85, highlighting its effectiveness in distinguishing between healthy and sick sheep. Despite these successes, the implementation faced challenges, including sensor integration, power management, and environmental noise, which were addressed through careful design and optimization. The field tests confirmed the system’s adaptability to real-world farming conditions, offering a practical tool for early disease detection that can enhance herd management and reduce economic losses for farmers. Ethically, the system prioritizes animal welfare by ensuring non-invasive monitoring and farmer privacy through secure data practices, aligning with the broader goals of sustainable and responsible smart farming. These findings underscore the system’s readiness for deployment while identifying areas for future improvement, such as enhanced power efficiency and broader dataset availability. Ultimately, this work lays a strong foundation for advancing livestock health monitoring, contributing to the future of precision agriculture and improved animal welfare.

# General Conclusion

This project has demonstrated the feasibility and potential of using a smart collar equipped with multimodal sensors and AI algorithms for the early detection of "*peste des petits ruminants (PPR)*" in sheep. By integrating temperature, heart rate, respiration, movement, and location data, the system can monitor animal health continuously and non invasively. Through feature engineering and model training, the project showed that deep learning models particularly the LSTM can effectively distinguish between healthy and infected animals based on sensor inputs.

The results confirm the initial hypotheses: intelligent collars can detect early symptoms of PPR with high accuracy, AI based classification significantly outperforms manual or threshold based methods, and the reduction in detection delay offers promising prospects for improved disease management and reduced outbreaks in flocks.

Beyond technical achievements, the system contributes to the broader vision of precision livestock farming, where real time monitoring enhances animal welfare and agricultural productivity. While some symptoms remain difficult to detect directly due to comfort or sensor limitations, the modular design of the system allows for future improvements such as new types of sensors or data fusion techniques.

This work lays the groundwork for real world deployment and further research, especially in scaling the solution to larger herds, validating the model with more diverse datasets, and integrating alerts into farm management platforms. Ultimately, such systems can revolutionize veterinary monitoring practices, leading to more resilient and efficient livestock farming.

# Bibliography

- [A<sup>+</sup>18] H. Ahmed et al. Dermatophytosis in small ruminants: Prevalence, species identification, and zoonotic concern. *Veterinary Microbiology*, 215:27–32, 2018.
- [A<sup>+</sup>21a] G. Alemayehu et al. Respiratory diseases in small ruminants in ethiopia. *Veterinary World*, 14(9):2383–2389, 2021.
- [A<sup>+</sup>21b] F. A. P. Alvarenga et al. Opportunities for precision livestock farming in ruminants. *Animal*, 15(S1):100204, 2021.
- [ABV<sup>+</sup>24] Muhammad Furqan Arshad, Giovanni Pietro Burrai, Antonio Varcasia, Maria Francesca Sini, Fahad Ahmed, Giovanni Lai, Marta Polinas, Elisabetta Antuofermo, Claudia Tamponi, Raffaella Cocco, Andrea Corda, and Maria Luisa Pinna Parpaglia. The groundbreaking impact of digitalization and artificial intelligence in sheep farming. *Research in Veterinary Science*, 170:105197, 2024.
- [AEM<sup>+</sup>20] A. Ait El Mekki et al. Malnutrition in small ruminants: Causes and prevention. *African Journal of Agricultural Research*, 15(3):455–460, 2020.
- [AER<sup>+</sup>18] K. A. Abd El-Razik et al. Epidemiological studies on respiratory diseases in small ruminants in egypt. *Veterinary World*, 11(6):810–815, 2018.
- [AK15] Mariette Awad and Rahul Khanna. Support vector machine. In *Efficient Learning Machines*, pages 70–82. ApressOpen, first edition, 2015.
- [ALC<sup>+</sup>24] Mateus Araujo, Paulo Leitão, Marina Castro, José Castro, and Miguel Bernuy. Development of an iot-based device for data col-

lection on sheep and goat herding in silvopastoral systems. *Sensors*, 24(17), 2024.

- [ALECRM<sup>+</sup>23] Carlos Alberto Aguilar-Lazcano, Ismael Edrein Espinosa-Curiel, Jorge Alberto Ríos-Martínez, Francisco Alejandro Madera-Ramírez, and Humberto Pérez-Espinosa. Machine learning-based sensor data fusion for animal monitoring: Scoping review. *Sensors*, 23(5732), 2023.
- [B<sup>+</sup>23] M. Belkadi et al. Sheep health behavior analysis in machine learning: A short comprehensive survey. *ArXiv Preprint*, 2023.
- [Ber20] Israr Berrim. Sentiment analysis in arabic tweets. Master’s thesis, University of Kasdi Merbah, Ouargla, Algeria, 2019/2020.
- [Bre94] Leo Breiman. Bagging predictors. Technical Report 421, University of California, Berkeley, California 94720, 1994. Accessed 11/03/2020.
- [ER<sup>+</sup>20] K. A. El-Razik et al. Stress biomarkers in sheep under intensive management system. *Egyptian Journal of Veterinary Sciences*, 51(2):269–276, 2020.
- [FKL<sup>+</sup>13] Ludwig Fahrmeir, Thomas Kneib, Stefan Lang, et al. *Regression: Models, Methods and Applications*. Springer-Verlag, Berlin Heidelberg, first edition, 2013.
- [G<sup>+</sup>20] A. Goel et al. Iot-based stress monitoring in livestock. *Computers and Electronics in Agriculture*, 176:105638, 2020.
- [GdRMB<sup>+</sup>24] Pedro Gonçalves, Maria do Rosário Marques, Ana Teresa Belo, António Monteiro, João Morais, Ivo Riegel, and Fernando Braz. Exploring the potential of machine learning algorithms associated with the use of inertial sensors for goat kidding detection. *Animals*, 14(938), 2024.
- [HMB<sup>+</sup>24] Charlotte Hollevoet, Evelyne Meyer, Claudia Bahr, Daniel Berckmans, and Annelies Van Nuffel. Goats on the move: Evaluating machine learning models for goat activity analysis using accelerometer data. *Animals*, 14(1977), 2024.
- [HUU<sup>+</sup>24] Rajib Kumar Halder, Mohammed Nasir Uddin, Md. Ashraf Uddin, Sunil Aryal, and Ansam Khraisat. Enhancing k-nearest neighbor



- algorithm: a comprehensive review and performance analysis of modifications. *Journal of Big Data*, 11:113, 2024.
- [Jam06] Mor James. Dartmouth artificial intelligence conference. the next fifty years. *AI Magazine*, 27(4):2006, 2006.
- [KBF<sup>+</sup>22] Zineddine Kouahla, Ala-Eddine Benrazek, Mohamed Amine Ferrag, Brahim Farou, Hamid Seridi, Muhammet Kurulay, Adeel Anjum, and Alia Asheralieva. A survey on big iot data indexing: Potential solutions, recent advancements, and open issues. *Future Internet*, 14(1):19, 2022.
- [KG09] J. Kaler and L. E. Green. Recognition of lameness and decisions to catch for inspection among sheep farmers and specialists in gb. *BMC Veterinary Research*, 5:41, 2009.
- [Kim16] Kwang Gi Kim. Book review: Deep learning. *Healthcare informatics research*, 22(4):351, 2016.
- [KVK<sup>+</sup>24] Yiannis Kiourrekis, Natalia G. C. Vasileiou, Eleni I. Katsarou, Daphne T. Lianou, Charalambia K. Michael, Sotiris Zikas, Angeliki I. Katsafadou, Maria V. Bourganou, Dimitra V. Liagka, Dimitris C. Chatzopoulos, and George C. Fthenakis. The use of machine learning to predict prevalence of subclinical mastitis in dairy sheep farms. *Animals*, 14(16):2295, 2024.
- [LBHH23] Djakhdjakha Lynda, Farou Brahim, Seridi Hamid, and Cissé Hamadoun. Towards a semantic structure for classifying iot agriculture sensor datasets: An approach based on machine learning and web semantic technologies. *Journal of King Saud University-Computer and Information Sciences*, 35:101700, 2023.
- [M<sup>+</sup>14] L. S. B. Mellau et al. Epidemiology of peste des petits ruminants (ppr) in africa. *Veterinary Journal*, 201:206–212, 2014.
- [Mah20] Batta Mahesh. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*.*[Internet]*, 9(1):381–386, 2020.
- [MCLL24] Ruiqin Ma, Runqing Chen, Buwen Liang, and Xinxing Li. A xgboost-based prediction method for meat sheep transport stress using wearable photoelectric sensors and infrared thermometry. *Sensors*, 24:7826, 2024.

- [MKH<sup>+</sup>24] Ghulam Mohyuddin, Muhammad Adnan Khan, Abdul Haseeb, Shahzadi Mahpara, Muhammad Waseem, and Ahmed Mohammed Saleh. Evaluation of machine learning approaches for precision farming in smart agriculture system: A comprehensive review. *IEEE Access*, 12:60155–60184, 2024.
- [Nil98] Nils J. Nilsson. *Artificial Intelligence: A New Synthesis*. Morgan Kaufmann Publishers, San Francisco, CA 94104-3205 USA, 1998. [Online]. 493p. Available at <https://books.google.dz/books?id=GYOFSd6fETgC&printsec=frontcover&dq=Artificial+Intelligence:+A+New+Synthesis&hl=fr&sa=X&ved=2ahUKEwiWlp3MvsvrAhVkxoUKHbKCBpoQuwUwAHoECAUQCg#v=onepage&q=Artificial%20Intelligence%3A%20A%20New%20Synthesis&f=false> (accessed 18/06/2020).
- [NKA<sup>+</sup>23] Muhammad Noor, Muhammad Khan, Hani Aljuaid, Hoda El-Sayed, Fahd Alghamdi, and Adeel Ur Rehman. Sheep health behavior analysis in machine learning: A short comprehensive survey. *Animal Artificial Intelligence*, 2:100194, 2023.
- [OIE22] OIE. Anthrax: Technical disease card. *World Organisation for Animal Health*, 2022.
- [Pav00] YU.L Pavlov. *Random Forest*. Ridderprint BV Ridderkerk, The Netherlands, first edition, 2000. Accessed 06/05/2024.
- [R<sup>+</sup>07] O. M. Radostits et al. *Veterinary medicine: A textbook of the diseases of cattle, sheep, pigs, goats and horses*. 2007.
- [R<sup>+</sup>18] A. Rashid et al. Pasteurellosis in sheep and goats: A review. *Veterinary Research International*, 6(3):77–82, 2018.
- [RGBR20] L. Ruiz-Garcia, P. Barreiro, and J. I. Robla. Wearable sensor systems for animal monitoring: A review. *Computers and Electronics in Agriculture*, 69(1):1–15, 2020.
- [RRH<sup>+</sup>24] Ali Rohan, Muhammad Saad Razaq, Md. Junayed Hasan, Furqan Asghar, Ali Ashif Bashir, and Tania Dottorini. Deep learning applications for livestock behavior recognition: A systematic literature review. *Computers and Electronics in Agriculture*, 224:109115, 2024.

- [SMS20] K Sindhu Meena and S Suriya. A survey on supervised and unsupervised learning techniques. In *Proceedings of international conference on artificial intelligence, smart grid and smart city applications: AISGSC 2019*, pages 627–644. Springer, 2020.
- [SS16] Shan Suthaharan and Shan Suthaharan. Decision tree learning. *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*, pages 237–269, 2016.
- [TSH<sup>+</sup>23] Kirk E. Turner, Ferdous Sohel, Ian Harris, Mark Ferguson, and Andrew Thompson. Lambing event detection using deep learning from accelerometer data. *Computers and Electronics in Agriculture*, 208:107787, 2023.
- [TTH<sup>+</sup>23] Kirk E. Turner, Andrew Thompson, Ian Harris, Mark Ferguson, and Ferdous Sohel. Deep learning based classification of sheep behaviour from accelerometer data with imbalance. *Information Processing in Agriculture*, 10(3):377–390, 2023.
- [Wei05] Sanford Weisberg. *Applied Linear Regression*. John Wiley & Sons, Hoboken, New Jersey, third edition, 2005.
- [XDW<sup>+</sup>25] Shun Xiao, Niranjana K. Dhand, Zhiyong Wang, Kun Hu, Peter C. Thomson, John K. House, and Mehar S. Khatkar. Review of applications of deep learning in veterinary diagnostics and animal health. *Frontiers in Veterinary Science*, 12:1511522, 2025.
- [ZH<sup>+</sup>24] Hafidh Zemour, Ilyes Hadbaoui, Malika Zoubeidi, Abdelkader Berrani, Azzedine Mouhous, Fathi Abdellatif Belhouadjeb, Abdelkader Ammam, and Mohamed Sadoud. Economic and analysis of the sheep meat value chain in algeria. *Les Cahiers du Cread*, 40(2):93–121, 2024.
- [ZZW<sup>+</sup>18] Y. Zhou, J. Zhang, S. Wang, Y. Chen, and S. Xu. Classification of eating behaviours in sheep using wearable sensors. *Sensors*, 18(10):3532, 2018.
- [ZZZ22] Y. Zhou, Y. Zhang, and Y. Zhang. A review of smart livestock farming technologies: Applications, challenges, and future trends. *Sensors*, 22(3):1070, 2022.

# Webographie

- [W1], Traditional breeding in Algeria. <https://news.cgtn.com/news/2023-07-18/Algerians-flock-to-sheep-market-ahead-of-holidays-1lxooj7LRaE/index.html>, Last access : 14/05/2025.
- [W2], internet of things applications [https://www.researchgate.net/figure/Examples-of-internet-of-things-applications-79\\_fig5\\_327227837](https://www.researchgate.net/figure/Examples-of-internet-of-things-applications-79_fig5_327227837), Last access : 14/05/2025.
- [W4] , [https://jmarwane.github.io/VTECL/img/ml\\_types.png](https://jmarwane.github.io/VTECL/img/ml_types.png), Last access : 11/05/2025.
- [W5], [https://cdn.prod.website-files.com/614c82ed388d53640613982e/63ef769f6a877d715fa75008\\_supervised%20vs%20Unsupervised%20learning.jpg](https://cdn.prod.website-files.com/614c82ed388d53640613982e/63ef769f6a877d715fa75008_supervised%20vs%20Unsupervised%20learning.jpg), Last access : 11/05/2025.
- [W6] , <https://www.datacamp.com/blog/supervised-machine-learning>, Last access : 11/05/2025.
- [W7] , <https://www.ibm.com/fr-fr/think/topics/random-forest>, Last access : 7/04/2025.
- [W8], <https://www.geeksforgeeks.org/random-forest-algorithm-in-machine-learning/>, Last access : 12/05/2025.
- [W9], <https://www.ibm.com/think/topics/support-vector-machine>, Last access : 13/04/2025.
- [W10], <https://spotintelligence.com/2024/05/06/support-vector-machines-svm/>, Last access : 14/05/2025.
- [W11], <https://www.geeksforgeeks.org/k-nearest-neighbours/>, Last access : 14/05/2025.
- [W12] , <https://www.ibm.com/think/topics/decision-trees>, Last access : 13/04/2025.
- [W13], <https://www.appliedaicourse.com/blog/decision-tree-in-machine-learning/>, Last access : 14/05/2025.

- [W14], <https://www.ibm.com/think/topics/xgboost>, Last access : 13/04/2025.
- [W15], [https://en.wikipedia.org/wiki/Gradient\\_boosting](https://en.wikipedia.org/wiki/Gradient_boosting), Last access : 13/04/2025 a 03:49.
- [W16], <http://thedata scientist.hashnode.dev/how-to-use-gradient-boostered-decision-trees>, Last access : 14/05/2025.
- [W17], <https://pub.towardsai.net/gradient-boosting-technique-b3dbb7069b74>, Last access : 14/05/2025.
- [W18], <https://www.ibm.com/think/topics/deep-learning>, Last access : 25/03/2025 a 04:10.
- [W19], <https://getthematic.com/insights/what-is-deep-learning/>, Last access : 14/05/2025.
- [W20], <https://www.ionos.co.uk/digitalguide/websites/web-development/convolutional-neural-networks/>, Last access : 14/05/2025.
- [W24], <https://medium.com/@RobuRishabh/understanding-and-implementing-faster-rcnn>, Last access : 25/03/2025 a 03:59.
- [W25], <https://docs.ultralytics.com/fr/#where-to-start>, Last access : 25/03/2025 a 03:10.
- [W26], <https://www.ibm.com/fr-fr/think/topics/recurrent-neural-networks>, Last access : 25/03/2025 a 04:10.
- [W27], [https://en.wikipedia.org/wiki/Multilayer\\_perceptron](https://en.wikipedia.org/wiki/Multilayer_perceptron), Last access : 25/03/2025 a 03:49.
- [W28], <https://datascientest.com/long-short-term-memory-tout-savoir>, Last access : 15/05/2025.
- [W29], <https://www.sciencedirect.com/topics/computer-science/multilayer-perceptron>, Last access : 15/05/2025.
- [W30], <https://www.smartshepherd.com.au/>, Last access : 18/04/2025 a 23:57.
- [W31], <https://www.nzherald.co.nz/the-country/news/smart-collar-a-game-changer-for-sheep-farmers/CQ23NPDG63UPDFLF5YKADZEWD/>, Last access : 18/04/2025 a 23:57.
- [W32], <https://www.zdnet.com/article/internet-of-sheep-why-worlds-biggest-nb-1000000-sheep-are-getting-smarter/>, Last access : 18/04/2025 a 23:00.
- [W33], <https://digitanimal.co.uk/product/digitanimal-gps-sheep-tracker/>, Last access : 18/04/2025 a 00:50.
- [W34], <https://digitanimal.co.uk/product/digitanimal-gps-sheep-tracker/>, Last access : 18/04/2025 a 00:00.
- [W35], <https://cerestag.com/blogs/press/sheep-tags-catch-mock-sheep-rustlers>, Last access : 18/04/2025 .

- [W36], <https://learn.sparkfun.com/tutorials/esp32-lora-1-ch-gateway-lorawan-an-introduction>, Last access : 16/05/2025 .
- [W37], <https://www.amazon.in/Techtonics-MAX30102-Oximeter-Sensor-Module/dp/B0C5XZHDWH>, Last access : 16/04/2025 .
- [W38], <https://mhtronic.com/produit/capteur-de-vibration-sw-420/>, Last access : 16/04/2025 .
- [W39], <https://pmdway.com/products/neo-6m-gps-module>, Last access : 16/04/2025 .
- [W40], <https://askelectronics.co.ke/product/ds1820-stainless-steel-package-wat> Last access : 16/04/2025 .
- [W41], <https://www.rechargeable-batterypack.com/sale-13751093-3-7-v-150mah-lip.html>, Last access : 16/04/2025.
- [W42] <https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-t> Last access : 24/05/2025.
- [W43] <https://www.tutorialspoint.com/xgboost/xgboost-architecture.htm> Last access : 25/05/2025.
- [W44] [https://www.researchgate.net/figure/General-architecture-of-a-support-ve\\_fig3\\_348745187](https://www.researchgate.net/figure/General-architecture-of-a-support-ve_fig3_348745187) Last access : 25/05/2025.
- [W45] [https://www.researchgate.net/figure/The-architecture-of-the-LightGBM-met\\_fig3\\_384512974](https://www.researchgate.net/figure/The-architecture-of-the-LightGBM-met_fig3_384512974) Last access : 25/05/2025.
- [W46] <https://medium.com/@anishnama20/understanding-gated-recurrent-unit-gru-i> Last access : 25/05/2025.
- [W47] <https://numpy.org/devdocs//user/whatisnumpy.html> Last access : 3/06/2025.
- [W48] [https://www.w3schools.com/python/pandas/pandas\\_intro.asp](https://www.w3schools.com/python/pandas/pandas_intro.asp) Last access : 3/06/2025.
- [W49] <https://scikit-learn.org/stable/modules/preprocessing.html> Last access : 3/06/2025.
- [W50] [https://matplotlib.org/3.5.3/api/\\_as\\_gen/matplotlib.pyplot.html#module-matplotlib.pyplot](https://matplotlib.org/3.5.3/api/_as_gen/matplotlib.pyplot.html#module-matplotlib.pyplot) Last access : 3/06/2025.
- [W51] [https://www.tensorflow.org/api\\_docs/python/tf/keras/Model](https://www.tensorflow.org/api_docs/python/tf/keras/Model) Last access : 3/06/2025.
- [W52] [https://www.tensorflow.org/api\\_docs/python/tf/keras/Model](https://www.tensorflow.org/api_docs/python/tf/keras/Model) Last access : 3/06/2025.
- [W53] [https://scikit-learn.org/1.4/modules/model\\_evaluation.html#:~:text=Classification%20metrics-,The%20sklearn.,values%2C%20or%20binary%20decisions%20values](https://scikit-learn.org/1.4/modules/model_evaluation.html#:~:text=Classification%20metrics-,The%20sklearn.,values%2C%20or%20binary%20decisions%20values). Last access : 3/06/2025.

- [W54] [https://imbalanced-learn.org/stable/references/generated/imblearn.over\\_sampling.SMOTE.html](https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.SMOTE.html) Last access : 3/06/2025.
- [W55] [https://www.tensorflow.org/api\\_docs/python/tf/keras/layers](https://www.tensorflow.org/api_docs/python/tf/keras/layers) Last access : 3/06/2025.
- [W56] [https://www.tensorflow.org/api\\_docs/python/tf/keras/callbacks/Callback](https://www.tensorflow.org/api_docs/python/tf/keras/callbacks/Callback) Last access : 3/06/2025.
- [W57] <https://seaborn.pydata.org/> Last access : 3/06/2025.
- [W58] <https://xgboost.readthedocs.io/en/stable/> Last access : 3/06/2025.
- [W59] <https://lightgbm.readthedocs.io/en/stable/> Last access : 3/06/2025.
- [W60] <https://scikit-learn.org/stable/api/sklearn.svm.html> Last access : 3/06/2025.
- [W61] <https://scikit-learn.org/stable/inspection.html> Last access : 3/06/2025.
- [W62] <https://docs.arduino.cc/libraries/wifi/> Last access : 3/06/2025.
- [W63] <https://github.com/amcewen/HttpClient> Last access : 3/06/2025.
- [W64] <https://docs.arduino.cc/language-reference/en/functions/communication/wire/t> Last access : 3/06/2025.
- [W65] <https://docs.arduino.cc/libraries/onewire/> Last access : 3/06/2025.
- [W66] <https://docs.arduino.cc/libraries/dallastemperature/> Last access : 3/06/2025.
- [W67] <https://www.elbilad.net/national/%D8%A7%D9%84%D8%AC%D8%B2%D8%A7%D8%A6%D8%B1-%D8%AA%D8%AD%D8%B5%D9%8A-17-%D9%85%D9%84%D9%8A%D9%88%D9%86-%D8%B1%D8%A3%D8%B3-%D8%BA%D9%86%D9%85-120955> Last access : 16/06/2025.
- [W68] <https://www.annahar.com/arabic/politics/arabi-world/almaghreb-alarabi/213185/%D8%A7%D9%84%D8%AC%D8%B2%D8%A7%D8%A6%D8%B1-%D8%AA%D8%B9%D9%8A%D8%AF-%D8%A5%D8%AD%D8%B5%D8%A7-%D8%AB%D8%B1%D9%88%D8%AA%D9%87%D8%A7-%D8%A7%D9%84%D8%AD%D9%8A%D9%88%D8%A7%D9%86%D9%8A%D8%A9-%D9%88%D8%A7%D9%84%D8%B2%D8%B1%D8%A7%D8%B9%D9%8A%D8%A9-%D8%A8%D8%B9%D8%AF-%D8%AA%D9%81%D8%A7%D9%88%D8%AA-%D9%83%D8%A8%D9%8A%D8%B1-%D9%81%D9%8A-%D8%A7%D9%84%D8%A3%D8%B1%D9%82%D8%A7%D9%85> Last access : 16/06/2025.

# Start-up Annex

## Project Presentation

### The project idea (proposed solution)

- The business area of the project is the service and agri-tech field, with a focus on smart livestock monitoring.
- The genesis of the idea for the project stems from realizing the difficulties faced by farmers in detecting early signs of diseases like "*peste des petits ruminants (PPR)*"
- This is done by designing a smart collar that integrates sensors for temperature, heart rate, vibration, and GPS to continuously monitor the animal's health and behavior.
- The collected data is transmitted via LoRa technology or wifi and analyzed to detect symptoms such as fever, coughing, respiratory distress, or isolation from the group.
- Farmers receive real-time alerts when abnormal signs are detected, allowing for quick intervention and disease control.

### The proposed values

- Farmers gain from intelligent illness identification, expanded monitoring for other animals, and up-to-date, real-time data thanks to the smart collar.
- Continuously updated and real-time data from sensors, providing instant insight into the animal's health and behavior.
- Early detection of disease symptoms to enable prompt intervention and reduce animal mortality and financial losses.
- Providing actionable insights to farmers based on the analysis of collected health and behavioral data.



- Automatic alerts sent via SMS to inform farmers immediately when abnormal health indicators are detected.

## **Work team**

**kolli marwa :** Has skills in designing intelligent livestock monitoring systems, including smart collars for early disease detection in sheep, with expertise in web services, artificial intelligence, and sensor data analysis.

**Messahel ayoub :** Has skills in designing intelligent livestock monitoring systems, including smart collars for early disease detection in sheep, with expertise in web services, artificial intelligence, and sensor data analysis.

**Farou Brahim:** has skills in the field of IT and project management.

- Messahel Ayoub and Kolli Marwa are responsible for the development of a smart collar designed to detect Peste des Petits Ruminants (PPR) in sheep. Their tasks include the conception and design of the system, development of the web service, integration of sensors, data acquisition, and hardware testing. Additionally, users conduct testing of the collar on the sheep to validate its performance.
- Farou Brahim's role: Lead the project.

## **The project's objectives**

**Short term:** Develop and deploy a functional prototype of the smart collar for early detection of PPR in sheep and initiate field testing with local shepherds.

**Medium term:** Improve the collar's performance and reliability by integrating sensor data analysis and artificial intelligence, and begin collaborations with Algerian livestock farms to monitor sheep health on a larger scale.

**Long term:** Establish the smart collar as a leading solution for PPR detection across the Algerian sheep farming sector, achieving widespread adoption, contributing to disease control, and expanding the technology to other regions and related livestock diseases within five years. Additionally, extend the collar's capabilities to detect other diseases, provide tailored veterinary recommendations (such as suitable medications and vaccines), and adapt the solution for use with other animals, such as cattle.

## Project completion schedule

| Phase                                                                             | 1m | 2m | 3m | 4m | 5m | 6m |
|-----------------------------------------------------------------------------------|----|----|----|----|----|----|
| Preliminary studies:<br>Analysis of the local market,<br>Choice of business model | *  | *  |    |    |    |    |
| Service development                                                               |    |    | *  | *  | *  |    |
| Testing and launch                                                                |    |    |    |    | *  |    |
| Marketing and promotion                                                           |    |    |    |    |    | *  |

Table 1.5 – Project Completion Schedule

## Innovative Aspects

- **Real-time Early Disease Detection:** The collar continuously monitors vital health indicators (temperature, heart rate, respiratory rate, vibration for cough detection) to detect early signs of "*peste des petits ruminants (PPR)*" before clinical symptoms become obvious.
- **Multimodal Sensor Integration:** Combining multiple sensors (temperature, GPS, vibration, pulse oximetry, accelerometer) provides a comprehensive health profile, improving detection accuracy compared to single-sensor devices.
- **Remote Monitoring via Web Services:** Health data collected by the collar is transmitted to a web platform, allowing farmers and veterinarians to monitor sheep health remotely in real time, even in remote or rural areas.
- **User-friendly Testing and Feedback Loop:** The system includes an interface for users (farmers, vets) to provide feedback and validate alerts, improving the AI model continuously with real-world data.
- **Low-cost and Scalable Solution:** Designed to be affordable for small and medium-sized farms, the collar leverages cost-effective hardware (ESP32, widely available sensors) to facilitate mass adoption.
- **Adaptability to Harsh Environments:** The collar is rugged and designed to operate in challenging farm conditions (dust, rain, variable temperatures), ensuring reliable operation in real agricultural settings.
- **Integration with Existing Farm Management Practices:** The solution complements traditional livestock monitoring, offering a smart, data-driven approach without disrupting established workflows.
- **Potential for Expansion:** The platform and hardware can be adapted to detect other diseases or monitor other types of livestock, making it a versatile tool for

the agricultural sector.

## Strategic Market Analysis

### The Market Segment:

**The potential market:** Owners of sheep farms, veterinary professionals to monitor and detect early signs of disease (such as PPR) in sheep using smart and connected solutions.

### The target market (Segment):

- **Livestock Farmers & Breeders**  
Seeking to:
  - Enhance animal health
  - Reduce mortality rates
  - Detect diseases early (e.g., PPR)
- **Veterinary Services**  
Goals:
  - Modernize disease surveillance
  - Empower rural farming with IoT-based solutions

### Measurement of Competition Intensity:

- **Low Direct Competition in the Local Market:** Few or no existing solutions in Algeria and surrounding regions specifically targeting PPR detection in sheep using smart collars, offering a first-mover advantage.
- **Emerging Global Solutions:** International companies are developing smart livestock monitoring tools, but most focus on general health tracking or dairy cattle, not specifically PPR in sheep.
- **Barriers to Entry for New Competitors:** Technical complexity, need for expertise in AI, sensor integration, and livestock health knowledge make it difficult for new entrants to quickly replicate the solution.
- **Fragmented Traditional Practices:** Most local farms rely on manual observation, with limited adoption of digital tools, lowering current competitive pressure but requiring education and awareness efforts.
- **Opportunities for Strategic Partnerships:** Low competition creates room for partnerships with government bodies, cooperatives, and NGOs to quickly scale and gain credibility.

- **Potential Competition from AgriTech Startups:** As digital agriculture grows, startups could emerge with similar concepts. Speed of development, user adoption, and innovation will be key to staying ahead.

### **Marketing Strategy:**

- **Workforce Efficiency:** Use of an agile development process and clear task distribution (sensors, AI, web services, field testing) to maximize team efficiency and accelerate time to market.
- **Promotional Campaigns:** Awareness campaigns targeting farmers through social media, agricultural fairs, and cooperatives, along with on-site demonstrations to build trust and encourage adoption of the smart collar.

## **Production and organization plan**

### **The Production Process**

- Partnerships of e-commerce , e-commerce social media pages.
- Development of the service.
- Testing and deployment.
- Marketing launch.

### **Supply:**

Companies, social media pages, etc.

### **Employees:**

Our smart collar project for disease detection in sheep is expected to create approximately 30 to 40 job opportunities. These positions include roles in embedded systems development, artificial intelligence and data analysis, veterinary consultation, field deployment, mobile/web development, customer support, and marketing. The interdisciplinary nature of the project ensures employment across both technical and agricultural sectors.

### **Special Discussion:**

For us, key partnerships include collaborations with veterinary institutions, agricultural cooperatives, and livestock health authorities. These partnerships are essential as they provide access to real-world data, facilitate field testing, and help build trust

with farmers. By working closely with these stakeholders, we establish mutually beneficial relationships that support the effective deployment of the smart collar, improve early disease detection, and promote animal welfare within the livestock sector.

## **Financial plan**

### **Costs and Charges:**

The identification of all necessary costs and investments is essential when launching our smart collar project, as it involves both technological development and field deployment in agricultural settings. This includes basic costs, annual and fixed costs, and other operational expenses. Here are the main aspects to consider:

#### **A) Initial Costs:**

##### **Infrastructure:**

- Acquisition of development space or leasing a small lab or office for hardware assembly and administration.
- Purchase of equipment such as 3D printers, computers, testing tools, and electronics benches.
- Setup of power backup, secure storage for components, and IoT testing environment.

##### **Technology and Development:**

- Procurement of sensors (temperature, GPS, O<sub>2</sub>, vibration, heart rate) and microcontrollers (ESP32).
- Development of firmware, AI models, and mobile/web platforms.
- 3D design and printing of collar casing prototypes.

##### **Field Testing and Deployment:**

- Cost of deploying collars on sheep in partner farms.
- Transportation, veterinary supervision, and data collection logistics.
- Compensation for pilot testers and feedback collection activities.

##### **Cloud Services and Software:**

- Subscriptions to cloud platforms for real-time data transmission, hosting, and storage.
- Licenses for AI training tools and software development environments.

### Equipment:

- **Computers and Workstations:** Development workstations used for embedded system programming, sensor data analysis, AI model training, and administrative tasks.
- **IoT and Testing Equipment:** Electronic testing benches, oscilloscopes, soldering stations, and multimeters for hardware development and prototyping.
- **3D Printing and Prototyping Tools:** 3D printers and associated materials for producing the collar casing and enclosure parts for field testing.
- **Servers and Cloud Services:** Cloud infrastructure for hosting the data analytics platform, storing time-series sensor data, and providing access to monitoring dashboards.
- **Networking Equipment:** Routers, switches, and modules for establishing local networks and connecting LoRa gateways to the cloud.
- **Software Tools:** Licenses and subscriptions for development environments (e.g., Arduino IDE, Python), AI toolkits, and data processing platforms.
- **Backup and Security Systems:** External hard drives, cybersecurity solutions, and automated backups to secure sensitive animal health data and prevent system failure.
- **Office Supplies:** Essential items such as desks, chairs, notebooks, filing systems, and documentation tools for the administrative and technical teams.

### Technology:

- **Sensor Technologies:** Integration of advanced sensors for temperature, heart rate, blood oxygen levels, vibrations (cough detection), and GPS to monitor vital signs and behavior in real time.
- **Embedded Systems:** Use of ESP32 microcontrollers for data acquisition, processing, and wireless communication, ensuring low power consumption and efficient performance in the field.
- **Artificial Intelligence Models:** Implementation of machine learning algorithms to analyze physiological and behavioral data for early detection of PPR symptoms.
- **Cloud Infrastructure:** Real-time data transmission and storage through cloud platforms, enabling centralized monitoring, historical data tracking, and alert systems for veterinarians and farmers.
- **Integration Tools:** Use of LoRa and IoT communication protocols to ensure long-range data transmission from remote farm locations to cloud databases.

- **User Interfaces:** Development of intuitive web interfaces that allow users (farmers, vets) to visualize animal health status, receive alerts, and track historical trends.
- **Analytics and Reporting Tools:** Dashboards and data visualization tools to support decision-making, intervention planning, and herd health management based on aggregated insights.

## B) Prototype Development Cost Breakdown

To better understand the financial feasibility of our solution, we compiled a detailed cost table outlining all the components and materials used to build the functional prototype of the smart collar. This table includes sensors, microcontrollers, connectivity modules, and prototyping tools, highlighting the affordability and scalability of the system for future production.

| Component                                         | Quantity | Unit Price (DA) | Total Price (DA) |
|---------------------------------------------------|----------|-----------------|------------------|
| ESP32 Microcontroller                             | 2        | 1300            | 2600             |
| DHT22 Temperature Sensor                          | 2        | 700             | 1400             |
| MAX30102 Pulse Oximeter                           | 3        | 550             | 1650             |
| SW-420 Vibration Sensor                           | 3        | 200             | 600              |
| GPS Module (NEO-6M)                               | 2        | 1300            | 2600             |
| LoRa Module (SX1278)                              | 1350     | 1350            | 1350             |
| 3D Printed Case (Material)                        | 1        | 1500            | 1500             |
| Breadboard and Wires                              | 2 sets   | 500             | 1000             |
| Battery and Charger                               | 1        | 1000            | 1000             |
| Misc. Components (Resistors, Diodes, Transistors) | -        | -               | 2000             |
| <b>Total</b>                                      |          |                 | <b>13,700 DA</b> |

Table 1.6 – Prototype Development Cost Breakdown

## C) Operational Costs:

### Personnel:

- **Salaries of Employees:** This includes compensation for hardware and embedded system developers, AI/data scientists, field technicians, veterinarians, project managers, mobile/web developers, and customer support staff.
- **Continuing Training for Staff:** Ongoing training sessions, certifications, and workshops for staff on livestock disease detection technologies, AI model updates, IoT protocols, and agricultural best practices.
- **Cloud Services and Hosting:** Ongoing expenses for cloud storage, server hosting, and data bandwidth to ensure continuous monitoring and accessibility

of sheep health data.

- **Software Subscriptions:** Recurring costs for maintaining software tools used for sensor data analysis, AI model management, system monitoring, and user interfaces.
- **Data Acquisition:** Costs related to collecting and processing physiological and behavioral data from field trials, including expenses for veterinary supervision, transportation, and on-site setup.

#### **Marketing and Customer Service:**

- **Marketing Campaigns:** Includes ongoing expenses for promoting the smart collar through agricultural fairs, social media campaigns, informational content targeting farmers and veterinarians, and partnerships with livestock cooperatives.
- **Customer Service Management:** Covers the costs of providing technical support to users (farmers and vets), responding to inquiries, managing product returns or malfunctions, and ensuring user satisfaction and proper collar usage.

#### **D) Other Costs:**

- **Liability Insurance:** Covers incidents related to the operation of the collars on livestock, including malfunction or damage.
- **Data Insurance:** Protects against potential loss or corruption of health monitoring data collected from the smart collars.
- **Licensing and Permits:** Costs associated with obtaining necessary regulatory approvals, veterinary certifications, and operational licenses.
- **Compliance with Regulations:** Ensuring adherence to animal welfare laws, data privacy regulations (for cloud data), and wireless communication standards in rural areas.

#### **E) Recurring Costs:**

- **Maintenance of Equipment and Devices:** Regular upkeep and calibration of smart collars, sensors, and base stations deployed in the field.
- **Renewal of Licenses and Permits:** Annual fees for renewing regulatory authorizations related to animal health technology and wireless communication.
- **System Updates and IT Maintenance:** Continuous software updates, firmware patches for the devices, and IT infrastructure maintenance including servers and databases.



## Methods and Sources of Obtaining Financing:

To finance this project, several methods and sources of funding can be explored:

### A) Internal Financing:

- **Use of the Company's Own Funds:** Initial capital provided by the founders or existing stakeholders to kickstart development and prototyping.
- **Reinvestment of Profits:** Profits generated from early sales or services reinvested into scaling production, marketing, or R&D.

### B) External Financing:

#### Bank Loans:

- **Long-term Loans:** To cover significant upfront investments such as mass production of collars, acquisition of advanced sensors, and infrastructure setup.
- **Lines of Credit:** To manage short-term operational costs including marketing, logistics, and maintenance.

#### Investors:

- We aim to attract investors interested in innovative agricultural technologies, particularly in smart livestock monitoring and early disease detection systems.
- Strategic partnerships with venture capitalists or stakeholders specializing in agri-tech, Internet of Things (IoT), and artificial intelligence will be actively pursued.
- Opportunities will be created to showcase the smart collar in agricultural fairs, innovation contests, and pilot programs to increase visibility and attract large-scale funding.

#### Subsidies and Aid:

- Explore government subsidies and grants available for innovative agricultural and veterinary technologies, particularly those focused on livestock health and rural development.
- Participate in start-up incubators and accelerators specialized in agri-tech, IoT, and artificial intelligence to benefit from mentorship, funding, and networking opportunities.

**Crowdfunding:** Utilize crowdfunding platforms to raise funds from a large number of small-scale contributors who are interested in supporting innovative agricultural solutions. This approach also helps increase public awareness of the project, create a community around the product, and validate interest from the market before scaling.

### C) Obtaining Reimbursement:

A detailed payment schedule can help plan the repayment of borrowed or invested funds. This table must include:

#### Scheduling Repayments:

- **Payment Schedule:**
  - Details of amounts to be repaid and deadlines for each source of financing.
  - Applicable grace periods and interest rates.
- **Cash Flow:**
  - Cash flow forecasting to ensure that the company has sufficient liquidity to meet repayment deadlines.
  - Adjustment of expenses and income to avoid deficits.

#### Payment Table:

| Date       | Amount         | Financing Type                  | Due Date | Remaining Balance |
|------------|----------------|---------------------------------|----------|-------------------|
| 2025-07-01 | +1,000,000 DZD | Initial Investment              | N/A      | 1,000,000 DZD     |
| 2025-08-15 | -500,000 DZD   | Sensor and Hardware Acquisition | N/A      | 500,000 DZD       |
| 2025-09-30 | +1,500,000 DZD | Additional Funding              | N/A      | 2,000,000 DZD     |
| 2025-09-30 | -750,000 DZD   | Prototype Development           | N/A      | 1,250,000 DZD     |
| 2025-10-15 | -300,000 DZD   | Field Testing on Sheep          | N/A      | 950,000 DZD       |
| 2025-11-30 | -400,000 DZD   | Data Analysis                   | N/A      | 550,000 DZD       |

The table above is a simplified example. It is essential to keep this payment table up to date and to rigorously monitor all key performance indicators related to the smart collar's deployment including disease detection accuracy, device durability, and user feedback. This ensures precise tracking of development stages, validates the effectiveness of the health-monitoring model in real conditions, and enables continuous improvement of the system. By maintaining detailed financial records and field data, the management and enhancement of the intelligent disease-detection

solution can be conducted efficiently, ensuring long-term impact and sustainability in the livestock sector.

# THE BUSINESS MODEL CANVAS

## KEY PARTNERS

- Minister of Agriculture and Rural Development.
- The Business Incubator at the University of Guelma
- Animal Welfare Associations
- Veterinary Clinics and Professionals
- Farmers' Associations
- Agricultural Supply Stores
- Fournisseurs IoT (capteurs, connectivity réseau).

## KEY ACTIVITIES

- Development of Smart Collar Technology
- Platform Development and Maintenance
- Data Collection and Analysis
- Customer Support and Issue Resolution
- Marketing

## KEY RESOURCES

- Smart Collar Technology & IoT Infrastructure
- Data Analytics & AI Algorithms
- Experienced Team
- Web platform

## VALUE PROPOSITION

- Health status tracking through a digital dashboard.
- Notification via SMS.
- Updated & real-time data.
- Early detection to minimize losses.
- Decision support based on collected data.

## CUSTOMER RELATIONS

- 24/7 dedicated phone support line
- Social media (Facebook)
- Online and on-site training (tutorial videos, illustrated guides).

## CHANNELS

- **Direct:**
  - Sales through agricultural cooperatives and animal feed stores.
- **Online:**
  - E-commerce platform (mobile payment for rural areas).

## CUSTOMER SEGMENTS

- Farmers (Livestock Owners)
- Veterinary Clinics

## COST STRUCTURE

- Production Costs
- Platform Development and Maintenance
- Marketing and Advertising
- Logistics and Distribution costs
  - Human resources salary
  - hardware costs (sensors ,collars)
- software costs

## REVENUE STREAM

- Direct Sales ( 5500 DA )
- Additional services: Annual maintenance (1000 DA), sensor replacement.
- Sale of anonymized data: To public institutions or researchers (long-term partnerships).