

République Algérienne Démocratique et Populaire.
Ministère de L'Enseignement Supérieure de la Recherche Scientifique.
Université 8 Mai 45 Guelma
Faculté des Mathématiques, d'Informatique et des Sciences de la Matière
Département d'Informatique



Mémoire de Fin d'Études Master

Filière : Informatique

Option : Systèmes Informatiques

Thème :

**Détection Automatique de l'Engagement Émotionnel dans un
Environnement d'Apprentissage Humain en Utilisant les
Techniques du Machine Learning**

présenté par :

SOUALA Akram Seyf Eddine

Devant Le Jury

M^{eme} BORDJIBA Yamina

Président

M^{eme} MEHENOUI Zohra

Encadreur

M^r KOUAHLA Mohamed Nadjib

Examineur

Soutenue le : 26/06/2023

Les émotions jouent un rôle central dans l'interaction et la compréhension humaine, et leur détection précise a un impact significatif dans des domaines tels que l'éducation, la santé mentale et les interfaces utilisateur. Des études ont montré que l'engagement émotionnel est crucial dans l'amélioration des résultats d'apprentissage dans les plateformes d'apprentissage en ligne.

Dans ce travail, nous proposons une approche pour la détection automatique de l'engagement émotionnel des apprenants dans les plateformes d'apprentissage en ligne. L'approche proposée est basée sur l'analyse des avis des apprenants pour pouvoir détecter leurs émotions. Les Machines à Vecteurs de Supports (SVM), la classification Naive Bayésienne (NB), et le modèle de Bidirectional Encoder Representations from Transformers (BERT) ont été utilisés dans la détection des émotions. Le modèle de BERT a montré une grande capacité dans la détection des émotions à partir du langage naturel.

Mots clés: Engagement émotionnel, BERT, SVM, NB, Langage naturel, Détection automatique.

Emotions play a central role in human interaction and understanding, and their accurate detection has a significant impact in fields such as education, mental health, and user interfaces. Studies have shown that emotional engagement is crucial in improving learning outcomes in online learning platforms.

In this work, we propose an approach for the automatic detection of learners' emotional engagement in online learning platforms. The proposed approach is based on analyzing learners' feedback to detect their emotions. Support Vector Machines (SVM), Naive Bayes classification (NB), and the Bidirectional Encoder Representations from Transformers (BERT) model have been used in emotion detection. The BERT model has shown great capability in detecting emotions from natural language.

Keywords: Emotional engagement, BERT, SVM, NB, Natural language, Automatic detection.

ملخص

العواطف تلعب دورًا أساسيًا في التفاعل وفهم التفاعل البشري، واكتشافها الدقيق له تأثير كبير في مجالات مثل التعليم والصحة العقلية وواجهات المستخدم. أظهرت الدراسات أن الارتباط العاطفي يعد أمرًا حاسمًا في تحسين نتائج التعلم في منصات التعلم عبر الإنترنت. في هذا العمل، نقترح نهجًا للكشف التلقائي عن الارتباط العاطفي للمتعلمين في منصات التعلم عبر الإنترنت. يستند النهج المقترح على تحليل آراء المتعلمين للكشف عن عواطفهم. تم استخدام أجهزة دعم القطع المرجحة (SVM) وتصنيف نايف بايز (NB) ونموذج الترميز المتوالي المزدوج من المحولات (BERT) في كشف العواطف.

أظهر نموذج BERT قدرة كبيرة في كشف العواطف من خلال اللغة الطبيعية. الكلمات المفتاحية: الارتباط العاطفي، BERT، SVM، NB، اللغة الطبيعية، الكشف التلقائي.

À ma grand-mère bien-aimée "Youma Daida", Ces mots sont une dédicace spéciale pour vous, la matriarche bienveillance de notre famille. Votre sagesse, votre amour inconditionnel et vos précieux conseils ont été des piliers tout au long de ma vie. Cette dédicace est dédiée tout particulièrement à vous, ma chère grand-mère, pour votre amour sans limites.

*À mon père "Salah" ma mère "Mounira" et ma sœur "Amira",
Votre soutien inébranlable et votre amour inconditionnel ont été les fondations de ma vie. Vous avez toujours été là pour moi, me guidant, m'inspirant et me donnant la force de poursuivre mes rêves. Cette dédicace est pour vous, ma famille, pour votre présence constante et votre soutien indéfectible.*

*À mes amis, "Nassima", "Zaki", "Sabar" et "Chilla" ,
Votre amitié a illuminé ma vie et rendu chaque moment inoubliable. Nassima, ta présence a apporté une énergie spéciale à notre cercle d'amis. Zaki et Sabar et Chilla vous avez été mes partenaires de confiance, mes frères de cœur, toujours prêts à partager des rires, des conseils et des aventures.*

REMERCIEMENTS

Je commence par exprimer ma plus profonde gratitude à Allah pour Ses innombrables bénédictions, Son guidance et Sa force tout au long de mon parcours académique.

Je suis également profondément reconnaissant envers ma superviseur "Mehenaoui Zohra", pour son précieux soutien, ses encouragements et ses critiques constructives tout au long de mon travail de recherche. Son expertise, ses conseils et son mentorat ont été essentiels pour façonner ma recherche et ma carrière académique.

J'adresse mes sincères remerciements aux membres du jury pour avoir pris le temps d'évaluer ma thèse et pour leurs précieux commentaires et suggestions. Je remercie également tous mes enseignants pour avoir transmis leurs connaissances et leur expertise, et pour m'avoir inspiré à poursuivre l'excellence académique.

Je suis profondément reconnaissant envers ma famille pour leur amour inconditionnel, leur soutien et leur motivation tout au long de mes études.

Enfin, je remercie mes amis pour leur encouragement constant, leur compréhension et leur soutien indéfectible.

Sans la contribution de chacune et chacun de ces individus, cette réussite n'aurait pas été possible. Je vous remercie du fond du cœur.

Merci.

Table des matières

RÉSUMÉ	I
Abstract	II
	III
DÉDICACE	IV
REMERCIEMENTS	V
Table des matières	VI
Liste des Figures	IX
INTRODUCTION GENERALE	1
CHAPTER I L'Engagement Émotionnel	3
I.1 Introduction	3
I.2 Définition de l'engagement	3
I.3 Type d'engagement	3
I.3.1 Emotionnel	4
I.3.2 Cognitive	4
I.3.3 Comportementale	4
I.4 Définition de l'émotion	5
I.5 Les applications de l'émotion	5
I.6 Sensors ou capteurs pour la reconnaissance de l'émotion	6
I.6.1 Capteurs visuels	6
I.6.2 Capteurs audio	7
I.6.3 Capteurs radar	8
I.6.4 Autres capteurs physiologiques	8
I.6.5 Fusion multi-capteurs	9
I.7 Méthode de détection de l'état émotionnel	10
I.7.1 Approches Machine Learning	10
I.7.1.1 Apprentissage supervisé	11

I.7.1.2	Apprentissage non supervisé	11
I.8	Approches deep learning	11
I.8.1	Modèle de transformateur	12
I.8.2	Modèle pré-entraîné de BERT	12
I.8.3	Approche basée sur le lexique	14
I.8.3.1	Approche manuelle	14
I.8.3.2	Approche basée sur le corpus	14
I.8.3.3	Approche basée sur les dictionnaires	14
I.8.4	Approche hybride	15
I.8.5	Autres Approches	15
I.9	Quelques travaux dans la prédiction des émotions	15
I.10	Conclusion	16
CHAPTER II Conception		18
II.1	Introduction	18
II.2	Objectif	18
II.3	Description de l'approche proposée	18
II.4	Prétraitement	19
II.4.1	Minuscule (Lowercase)	20
II.4.2	Caractères non alphabétiques(Non-alphabetic characters)	21
II.4.3	Mot vide (Stop words)	22
II.4.4	lemmatisation(Lemmatization)	23
II.4.5	Tokénisation(Tokenization)	24
II.5	Extraction de caractéristiques: Bag of Words (BOW)	26
II.6	Apprentissage	27
II.7	Décision	27
II.7.1	La classification Naive bayésienne (NB)	28
II.7.2	Les Machines à Vecteurs de Support	29
II.7.3	Bidirectional Encoder Representations from Transformers (BERT)	29
II.8	Conclusion	33
CHAPTER III Implémentation		34
III.1	Introduction	34
III.2	Environnement	34
III.2.1	Google Colaboratory	34
III.2.2	Apprentissage sur GPU	35
III.3	Langage de programmation et bibliothèque utilisée	35
III.3.1	Python	35
III.3.2	NLTK	36
III.3.3	TensorFlow	37

III.3.4 Numpy	37
III.4 La base de données (Data Set)	37
III.4.1 Description	37
III.4.2 Répartition de la base de données	37
III.5 Apprentissage et test	38
III.6 Les résultats obtenus	41
III.7 Comparaison entre l'approche de Machine Learning et l'approche de Deep Learning	42
III.8 Conclusion	42
CONCLUSION GENERAL	43

Liste des Figures

I.1	Processus de reconnaissance des expressions faciales.[1]	7
I.2	Le processus de SER.[1]	7
I.3	Processus de reconnaissance des émotions à l'aide d'un capteur radar.[1]	8
I.4	Signaux physiologiques détectés par un autre capteur physiologique.[1]	9
I.5	Fusion au niveau des caractéristiques.[1]	10
I.6	Approches de l'analyse des sentiments [2]	11
I.7	Architecture du modèle de transformateur.[3]	13
II.1	L'approche proposée.	19
II.2	Le prétraitement du Text (preprocessing)	20
II.3	miniscule	21
II.4	Non-alphabetic	22
II.5	Mot vide	23
II.6	lemmatisation	24
II.7	Tokénisation	25
II.8	Le prétraitemen du BERT.	26
II.9	Construction du vecteur de caracteristique.	28
II.10	Vecteur positive et negative adjective	28
II.11	Hyperplan de séparation	30
II.12	Pré-entraînement et fine-tuning BERT.[4]	31
II.13	BERT-large et bert-base.[5]	32
II.14	Architecture du modèle BERT.[4]	32
III.1	Interface de Colab.	35
III.2	Activation de GPU.	36
III.3	les dix premières lignes de la base de données.	38
III.4	Accuracy et Validation Accuracy sur Epoch.	40
III.5	Loss et Validation Loss sur Epochs.	40
III.6	Comparaison des résultats	42

Liste des Tableaux

I.1	Avantages et inconvénients des différents capteurs pour la reconnaissance des émotions.[1]	6
I.2	Quelques travaux connexes	16
II.1	Le schéma d’encodage de l’engagement émotionel.	29
III.1	La division de la base de données.	38
III.2	Les valeurs des paramètres du modèle	39
III.3	Résultats de la classification SVM	41
III.4	Résultats de la classification NB	42
III.5	Résultats de la classification BERT.	42

INTRODUCTION GENERALE

L'apprentissage en ligne devient de plus en plus l'option d'apprentissage la plus populaire. Comme l'affirment de nombreux chercheurs [6], l'émotion est considérée comme un élément essentiel de l'apprentissage en ligne. Il semble jouer un rôle essentiel au cœur de l'apprentissage, en influençant les processus cognitifs comme l'attention et la mémorisation des informations [7]. Les travaux antérieurs ont montré que la gestion et la régulation des émotions étaient étroitement liées à l'engagement émotionnel dans le contexte de l'apprentissage en ligne [8]. En d'autres termes, le manque de capacité à réguler les émotions peut entraîner de mauvaises conditions d'engagement émotionnel en ligne. L'engagement émotionnel fait référence aux émotions et aux intérêts des apprenants envers l'apprentissage et les autres participants. Contrairement à l'apprentissage en présentiel, l'absence de l'émotion de communication émotionnelle en face à face dans les environnements en ligne provoque l'isolement des apprenants, une mauvaise régulation émotionnelle et un manque d'engagement[9].

Des études importantes ont accordé une attention considérable à la prédiction de l'engagement émotionnel dans les environnements d'apprentissage en ligne.

L'objectif de ce travail est de proposer une approche pour la détection automatique de l'engagement émotionnelle dans un environnement d'apprentissage en ligne, en exploitant les avis des apprenants envers les ressources d'apprentissage présentés par leurs enseignants.

Trois techniques ont été exploitées pour la détection automatique des émotions des apprenants. Les deux premières issues des techniques de Machine Learning, à savoir: les Machines à Vecteurs de Supports (SVM) et la classification Naive Bayésienne (NB). La troisième technique est un modèle de Deep Learning orienté traitement automatique des langues, appelée BERT abréviation de Bidirectional Encoder Representations from Transformers.

Bert s'appuie sur la technologie des transformers pour ingérer des données séquentielles. Ce qui les rend particulièrement bien adaptés au traitement du langage naturel. L'innovation

clé de Bert consiste à mettre en œuvre l'entraînement bidirectionnel du transformer. Contrairement aux modèles de langue unidirectionnels, qui traitent le texte saisi de manière séquentielle, de gauche à droite ou de droite à gauche, l'encodeur mis en œuvre par Bert lit l'ensemble de la séquence de mots en une fois.

Pour valider l'approche proposée, des tests ont été faites sur la base de données collectée du site web Mark My Professor[10]. Cette base de données est constituée de 5200 avis d'apprenants concernant les cours présentés par leurs professeurs.

Le mémoire est organisé comme suit:

Le premier chapitre de ce mémoire présente un aperçu sur l'analyse des émotions et les techniques de détection des émotions.

Dans le deuxième chapitre, nous présentons la conception de notre système. nous décrivons le processus suivie par l'approche proposée, ainsi que les techniques de détection utilisées.

Le troisième chapitre est consacré à l'implémentation de notre système. Nous décrivons l'environnement de travail, ainsi que les bibliothèques utilisés. Les résultats obtenus sont présentés à la fin du chapitre.

Le mémoire est achevé par une conclusion générale et quelques perspectives.

I.1 Introduction

L'engagement des apprenants dans l'apprentissage en ligne est effectivement l'une des préoccupations majeures dans beaucoup de travaux de recherches [11]. L'engagement est un facteur essentiel du processus d'apprentissage, car il joue un rôle important dans les résultats scolaires et la motivation des élèves. Il a été démontré que l'engagement émotionnel, en particulier, a un impact sur la réussite scolaire, les compétences sociales . Avec la popularité croissante des environnements d'apprentissage en ligne, il est essentiel de comprendre et de mesurer l'engagement émotionnel des étudiants dans ces environnements.

I.2 Définition de l'engagement

L'engagement comme l'investissement psychologique et l'effort dirigés vers l'apprentissage, la compréhension, la maîtrise des connaissances, le développement des compétences ou des habiletés qui sont attendus [12]. Les recherches confirment que si les apprenants sont bien engagés, ils participeront plus activement aux activités d'apprentissage [11].

L'engagement envers l'apprentissage est lié au niveau de participation psychologique de la personne apprenante à l'activité et peut mieux prédire les réalisations en apprentissage.[13]

I.3 Type d'engagement

L'engagement d'apprentissage dans les discussions du forum est un concept multidimensionnel composé de trois composantes principales : l'engagement émotionnel, cognitif et comportemental.

I.3.1 Emotionnel

L'engagement émotionnel dans l'apprentissage en ligne implique la création d'un lien entre l'apprenant et le matériel d'apprentissage par l'évocation d'émotions [14]. Il se réfère aux réactions émotionnelles envers le contenu et le contexte d'apprentissage [14]. Cela peut être réalisé par diverses méthodes, telles que l'intégration de scénarios de la vie réelle, la narration, les jeux de rôle ou la création d'un lien personnel avec le matériel. En faisant appel aux émotions de l'apprenant, l'apprentissage en ligne peut accroître sa motivation, son attention et son intérêt, ce qui permet de mieux comprendre la matière et de mieux retenir les informations. De plus, l'engagement émotionnel peut également favoriser l'empathie et la compréhension du point de vue des autres, contribuant ainsi à l'efficacité globale de l'expérience d'apprentissage.

I.3.2 Cognitive

Dans l'apprentissage en ligne, l'engagement cognitif fait référence au niveau d'implication mentale active et d'attention qu'un apprenant porte au matériel d'apprentissage. Il implique de réfléchir et de traiter activement les informations présentées, plutôt que de simplement les recevoir passivement. L'engagement cognitif dans l'apprentissage en ligne peut être facilité par des activités interactives et stimulantes, telles que des tâches de résolution de problèmes, des simulations et des discussions. Il exige un effort de la part de l'apprenant et peut conduire à une compréhension plus profonde du matériel et à une meilleure rétention de l'information. En engageant les apprenants sur le plan cognitif, l'apprentissage en ligne peut aider à développer des compétences de pensée critique et encourager l'apprentissage indépendant.

I.3.3 Comportementale

Dans l'apprentissage en ligne, l'engagement comportemental fait référence au niveau de participation active et d'interaction d'un apprenant avec le matériel d'apprentissage. Ce type d'engagement implique de répondre physiquement au matériel par des actions telles que cliquer sur des liens, taper des réponses ou déplacer des objets dans des simulations. L'engagement comportemental dans l'apprentissage en ligne peut être facilité par des activités interactives qui demandent à l'apprenant de prendre des mesures en réponse au matériel, comme des quiz, des jeux ou des simulations interactives. En demandant aux apprenants de s'engager activement dans le matériel, l'apprentissage en ligne peut contribuer à améliorer leur compréhension du matériel et à accroître la rétention de l'information. De plus, l'engagement comportemental peut également accroître la motivation et l'intérêt pour le processus d'apprentissage.

I.4 Définition de l'émotion

L'émotion est l'expression globale de l'état physiologique et psychologique d'une personne, la reconnaissance des émotions a été proposée dans les années 1990s [15]

I.5 Les applications de l'émotion

Avec le développement rapide de la science et de la technologie, la reconnaissance des émotions est largement utilisée dans différents domaines tels que l'interaction homme-machine(IHM) [16] , le e-learning[17], la médecine [18], l'analyse psychologique [19], la surveillance de la sécurité [20], le pilotage intelligent [21], et l'industrie du divertissement [22].

La reconnaissance des émotions peut être obtenue en utilisant différentes méthodes de détection et différents capteurs. Les capteurs sont combinés avec des modèles algorithmiques avancés et des données riches pour former des systèmes d'interaction homme-ordinateur[23]. Dans le domaine du e-learning [24], la reconnaissance des émotions peut être utilisée pour détecter le statut d'apprentissage des apprenants et pour améliorer l'efficacité de l'apprentissage et prédire les résultats d'apprentissage. Dans le domaine de la médecine [25], les émotions peuvent être utilisées pour détecter l'état psychologique des patients et améliorer l'efficacité médicale. Dans le domaine de la psychanalyse [26], la reconnaissance des émotions de peut être utilisée pour aider à découvrir si une personne souffre d'autisme ou non. Dans les interrogatoires des suspects, la reconnaissance des émotions peut être utilisée dans le test d'authenticité [27]. Dans le domaine de la surveillance de sécurité, la reconnaissance des émotions peut être utilisée pour détecter l'état de somnolence du conducteur afin d'améliorer la sécurité de la conduite [28]. Cette technique peut également être utilisée pour identifier les émotions des personnes âgées, des nourrissons et des personnes atteintes de maladies particulières qui ne peuvent pas exprimer clairement leurs émotions [29]. La reconnaissance des émotions peut être obtenue en utilisant différentes méthodes de détection et différents capteurs (sensors). Les capteurs sont combinés avec des modèles algorithmiques avancés et des données riches pour former des systèmes d'interaction homme-ordinateur[23] ou des systèmes robots. Dans le domaine du e-learning [24], la reconnaissance des émotions peut être utilisée pour détecter le statut d'apprentissage des apprenants et pour améliorer l'efficacité de l'apprentissage et prédire les résultats d'apprentissage. Dans le domaine de la psychanalyse [26], il peut être utilisé pour aider à analyser si une personne souffre d'autisme. Cette technique peut également être utilisée pour identifier les émotions des personnes âgées, des nourrissons et des personnes atteintes de maladies particulières qui ne peuvent pas exprimer clairement leurs émotions [29].

I.6 Sensors ou capteurs pour la reconnaissance de l'émotion

Les capteurs utilisés pour la reconnaissance des émotions comprennent principalement des capteurs visuels, des capteurs audio, des capteurs radar et d'autres capteurs de signaux physiologiques qui peuvent collecter des signaux de différentes dimensions, et en même temps utiliser certains algorithmes pour réaliser une analyse émotionnelle. Différents capteurs ont différentes applications dans la reconnaissance des émotions. Le tableau I.1 répertorie les avantages et les inconvénients de différents capteurs pour la reconnaissance des émotions.

Capteurs	Avantages	Inconvénients
Capteurs visuels	Collecte de données simple grande évolutivité	Restreint par la lumière facile de causer des fuites de confidentialité.
Capteurs audio	Faible coût Large éventail d'applications	Manque de robustesse pour l'analyse de sentiments complexes
Capteurs radar	Surveillance à distance des signaux physiologiques	Le mouvement radial peut causer des perturbations
Capteurs texte	Grande quantité de données Traitement rapide	Précision limitée Dépendance du langage
Autres capteurs physiologiques	Possibilité de surveiller les signaux physiologiques représentant une émotion réelle	Invasif, nécessite d'être porté près de la surface de la peau.
Fusion multi-capteurs	Informations collectées plus riches plus grande robustesse	Les informations multicanaux doivent être synchronisées Le calcul de suivi est Relativement important

Table I.1: Avantages et inconvénients des différents capteurs pour la reconnaissance des émotions.[1]

I.6.1 Capteurs visuels

La reconnaissance des émotions basée sur des capteurs visuels est l'une des méthodes de reconnaissance des émotions les plus courantes. Elle présente l'avantage d'être peu coûteuse et permet une collecte de données simple. À l'heure actuelle, les capteurs visuels sont principalement utilisés pour la reconnaissance des expressions faciales (FER) [30] afin de détecter les émotions ou pour la technologie de photopléthysmographie à distance (rPPG) afin de détecter le rythme cardiaque [31]. La précision de ces méthodes diminue fortement

lorsque l'intensité lumineuse baisse.

Le processus de reconnaissance des expressions faciales est illustré dans la figure I.1. Les expressions faciales reflètent intuitivement les émotions des personnes. Il est difficile pour les machines de saisir les détails des expressions comme le font les humains [32]. Les expressions faciales sont faciles à dissimuler, ce qui entraîne des erreurs de reconnaissance des émotions [33]. Par exemple, dans certaines activités sociales, nous sourions poliment même si nous ne sommes pas d'humeur joyeuse [34].



Figure I.1: Processus de reconnaissance des expressions faciales.[1]

I.6.2 Capteurs audio

La langue est l'une des composantes les plus importantes de la culture humaine. Les gens peuvent s'exprimer ou communiquer avec les autres par le biais du langage. La reconnaissance vocale [35] a favorisé le développement de la reconnaissance des émotions vocales (SER) [36]. La parole humaine contient de nombreuses informations qui peuvent être utilisées pour la reconnaissance des émotions [37]. La compréhension de l'émotion dans l'information est essentielle pour que l'intelligence artificielle puisse s'engager dans un dialogue efficace. La SER peut être utilisée pour le dialogue dans les centres d'appel, les systèmes de réponse automatique, etc. [38]. La SER est complétée conjointement par l'extraction de caractéristiques acoustiques [39] et le marquage du langage [40]. Le processus de SER est illustré à la figure I.2

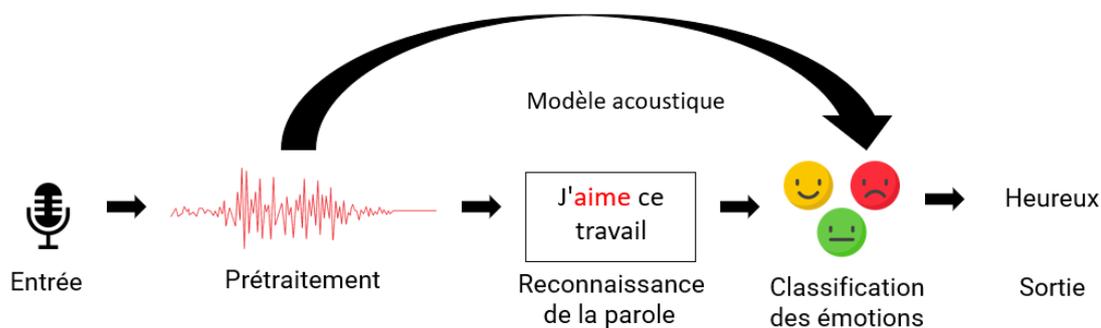


Figure I.2: Le processus de SER.[1]

I.6.3 Capteurs radar

Différentes émotions provoquent une série de réponses physiologiques, telles que des changements dans la fréquence respiratoire, la fréquence cardiaque, les ondes cérébrales, la pression artérielle, etc. Par exemple, l'excitation provoquée par le bonheur, la colère ou l'anxiété peut entraîner une augmentation du rythme cardiaque [41]. Les émotions positives peuvent augmenter la fréquence respiratoire, tandis que les émotions dépressives peuvent avoir tendance à inhiber la respiration [42]. La fréquence respiratoire affecte également la variabilité de la fréquence cardiaque (VFC), qui diminue à l'expiration et augmente à l'inspiration [43]. Actuellement, la technologie radar est largement utilisée pour la détection à distance des signes vitaux [44] et la détection sans fil [45]. Les capteurs radar peuvent utiliser le signal d'écho de la cible pour analyser les micro-mouvements de la poitrine causés par la respiration et les battements de cœur. Ils peuvent réaliser l'acquisition à distance de ces signaux physiologiques. Le processus global de reconnaissance des émotions à l'aide d'un capteur radar est illustré à la figure I.3

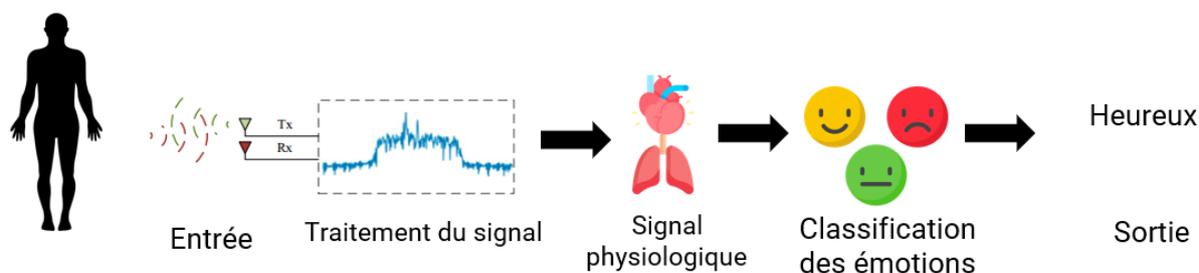


Figure I.3: Processus de reconnaissance des émotions à l'aide d'un capteur radar.[1]

I.6.4 Autres capteurs physiologiques

Il a été démontré que les émotions sont biologiques depuis les temps anciens. Des émotions excessives peuvent affecter le fonctionnement des organes vitaux. Aristote croyait que l'influence des émotions sur la physiologie se reflétait dans les changements d'états physiologiques, tels que l'augmentation du rythme cardiaque, l'augmentation de la température corporelle ou la perte d'appétit. William James a été le premier à proposer une théorie de la physiologie de l'émotion [46]. Il croyait que les stimuli externes déclenchent une activité dans le système nerveux autonome et produisent une réponse physiologique dans le cerveau. Par exemple, nous rions quand nous sommes heureux, nous rions quand nous sommes heureux, nous rampons quand nous avons peur et pleurons quand nous sommes tristes.

L'émotion humaine est un état mental spontané qui se reflète dans les changements physiologiques du corps humain et a un impact significatif sur notre conscience [47]. De nombreux autres signaux physiologiques du corps humain, tels que l'électroencéphalogramme

(EEG) [48], l'électrocardiogramme (ECG) [49], l'électromyographie (EMG) [50], la réponse galvanique cutanée (GSR) [51], le pouls du volume sanguin (BVP) [52] et électrooculogramme (EOG)[53], comme indiqué dans la figure I.4

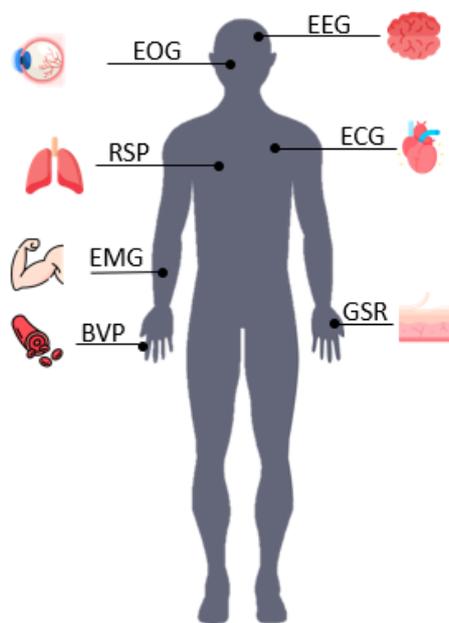


Figure I.4: Signaux physiologiques détectés par un autre capteur physiologique.[1]

I.6.5 Fusion multi-capteurs

L'interprétation unimodale des émotions présente certaines lacunes et l'identification générale précise des émotions complexes n'est pas possible. L'initiation affective multimodale fait référence à l'utilisation de signaux recueillis auprès d'observateurs supplémentaires pour atteindre et maintenir les résultats de l'initiation et les résultats de l'initiation. Une approche multimodale facilite le développement de l'intelligence émotionnelle. L'intelligence émotionnelle multimodale permet d'obtenir les meilleures performances en matière d'intelligence, mais augmente également la complexité de l'information en raison du nombre élevé de canaux. Un matériau trop exigeant pour une collection de données multimodales. La clarification multimodale des émotions est appliquée à diverses stratégies de fusion, qui sont principalement divisées en fusion au niveau du pixel, fusion au niveau du caractère et fusion au niveau de la décision.

La fusion au niveau du pixel [54] fait référence à la fusion directement à partir de l'original ; les informations sémantiques et la génération de signaux sont superposées, ce qui affecte l'effet de classification après la fusion. La période de travail de l'entreprise est toujours pauvre en informations redondantes. Le processus de fusion au niveau des caractéristiques

[55] est illustré à la figure I.5. Le processus de fusion au niveau des caractéristiques a lieu au cours de la première étape du processus de fusion. Les caractéristiques des différents signaux d'entrée sont extraites et combinées avec des caractéristiques à haute dimension. Enfin, les résultats sont envoyés au classificateur. La fusion au niveau des caractéristiques préserve une grande quantité d'informations importantes, ce qui permet d'examiner attentivement l'information. Mais en raison d'une quantité insuffisante de données ou d'un nombre insuffisant de détails, l'exigence de précision est finalement réduite.

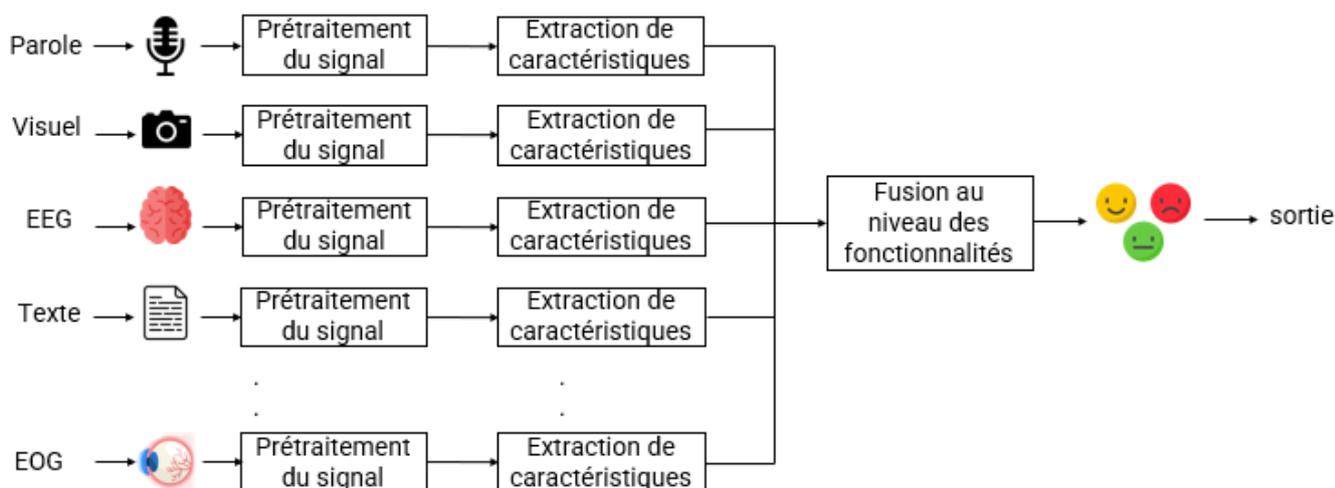


Figure I.5: Fusion au niveau des caractéristiques.[1]

I.7 Méthode de détection de l'état émotionnel

Les techniques d'analyse des sentiments peuvent être classées en cinq grandes approches : l'apprentissage automatique, l'apprentissage profond, l'approche basée sur le lexique, l'approche hybride et d'autres approches, comme le montre la figure I.6.

I.7.1 Approches Machine Learning

Les méthodes d'apprentissage automatique sont largement utilisées pour produire des modèles de classification des sentiments dans le domaine de l'analyse des sentiments. Ces méthodes construisent d'abord un ensemble d'apprentissage et étiquettent les données d'apprentissage en fonction du sentiment. Un ensemble de caractéristiques est ensuite extrait des données d'apprentissage et transmis à un modèle de classification tel que Naive Bayes (NB), Support Vector Machine (SVM), Logistic regression (LR), Random Forest (RF), etc. Après l'entraînement avec les

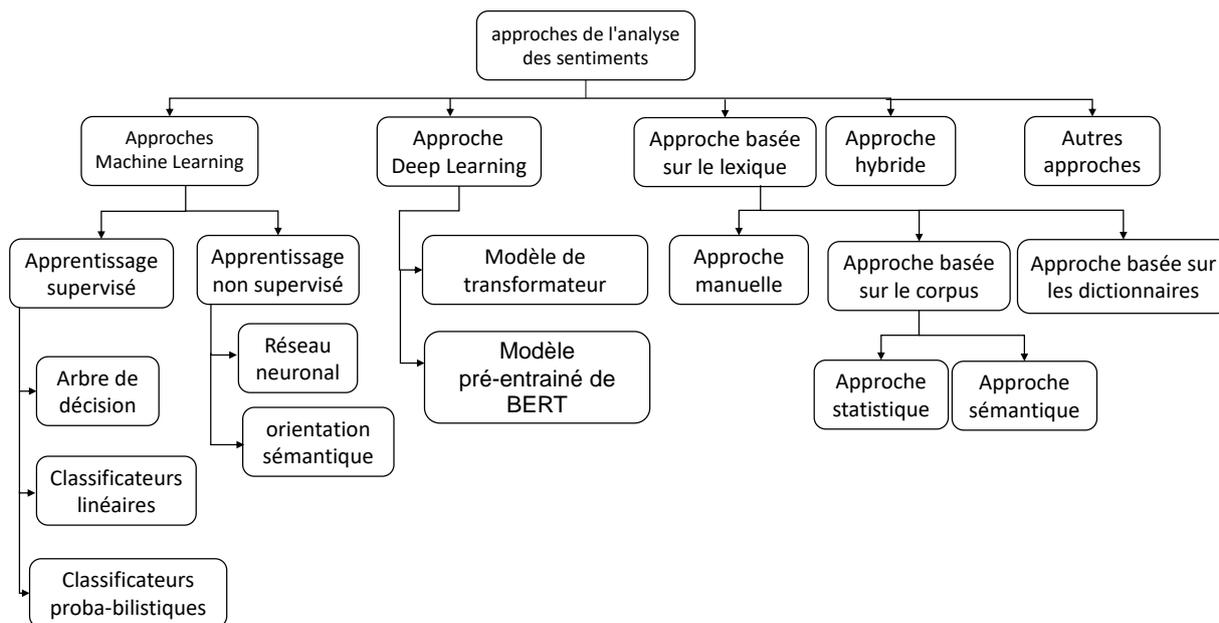


Figure I.6: Approches de l'analyse des sentiments [2]

étiquettes de sentiment, le classificateur peut être utilisé pour prédire l'orientation du sentiment d'un échantillon qui n'est pas annoté.[56]

I.7.1.1 Apprentissage supervisé

Les approches d'apprentissage supervisé utilisent des documents de formation étiquetés. L'apprentissage supervisé est basé sur la classification automatique de textes.[57]

I.7.1.2 Apprentissage non supervisé

Contrairement aux approches d'apprentissage supervisé, les approches d'apprentissage non supervisé ne dépendent pas du domaine et du sujet des données de formation. Les approches d'apprentissage non supervisé surmontent la difficulté de collecter et de créer des données d'entraînement étiquetées.[57]

I.8 Approches deep learning

Ces dernières années, l'apprentissage en profondeur a attiré beaucoup d'attention en raison de sa capacité à apprendre et à extraire automatiquement des caractéristiques significatives à partir de données brutes. Dans la détection des émotions, les techniques d'apprentissage en profondeur ont été appliquées avec succès à l'analyse de diverses modalités telles que le texte, la parole et les expressions faciales. Voici quelques méthodes couramment utilisées : les réseaux de neurones convolutifs (CNN), les réseaux de neurones récurrents (RNN),

l'apprentissage profond multimodal, l'apprentissage par transfert, les mécanismes d'attention et les réseaux antagonistes génératifs (GAN).

I.8.1 Modèle de transformateur

Le modèle Transformer utilise un mécanisme d'auto-attention au lieu des couches récurrentes ou convolutives traditionnelles. Ce mécanisme d'auto-attention permet au modèle d'apprendre la relation entre les éléments de la séquence d'entrée (langage naturel), ce qui en fait une innovation clé dans le domaine de l'apprentissage en profondeur pour traiter des séquences de données avec plus de parallélisme, conduisant à des systèmes pré-formés, comme BERT et ses variantes, T5, GPT, ces systèmes sont formés à l'aide de grands ensembles de données linguistiques comme le corpus Wikipedia et peuvent être affinés pour des tâches spécifiques. Un ensemble d'états cachés décrivant les dépendances entre les éléments de la séquence d'entrée est produit par le transformateur après que la séquence d'entrée est transformée en un espace de grande dimension et traitée à travers plusieurs couches d'auto-attention. Ces états cachés sont ensuite utilisés pour faire des prédictions sur des tâches objectives, telles que la traduction automatique ou la classification de texte. Les résultats des travaux des auteurs [3] montrent que Transformer surpasse les modèles de pointe précédents utilisant des réseaux de neurones récurrents et convolutifs. De plus, les auteurs mettent l'accent sur l'efficacité et la parallélisabilité de Transformer, ce qui facilite son apprentissage par rapport aux modèles précédents.

Les travaux de Vaswani et al [3] sont considérés comme un tournant dans le deep learning, ayant un impact majeur sur la création de modèles NLP. Ses concepts de base ont été largement utilisés et développés dans de nombreux travaux ultérieurs. la figure II.14 indique Architecture du modèle de transformateur

I.8.2 Modèle pré-entraîné de BERT

Jacob Devlin et al. [58] a proposé une nouvelle approche de pré-formation au traitement du langage naturel (NLP) appelée BERT (Représentations d'encodeurs bidirectionnels à partir de transformateurs). L'idée est de former un modèle profond de représentations bidirectionnelles sur un grand corpus de texte non étiqueté, tel que le corpus Wikipedia, puis de l'améliorer pour une tâche NLP spécifique. Parlons des principaux points du BERT :

Architecture : BERT est une architecture de réseau neuronal basée sur un transformateur similaire à celle introduite par Vaswani et al [3]. Cependant, BERT est bidirectionnel, ce qui signifie qu'il peut prendre en compte l'ensemble du contexte des mots dans une phrase, plutôt que le premier mot. Le contexte entier d'un mot dans une phrase, pas les mots avant ou après. avant ou après. BERT utilise un transformateur multicouche bidirectionnel pour générer une série d'incorporations de mots contextualisées. Il introduit également deux nouvelles tâches d'entraînement : la modélisation du langage masqué et

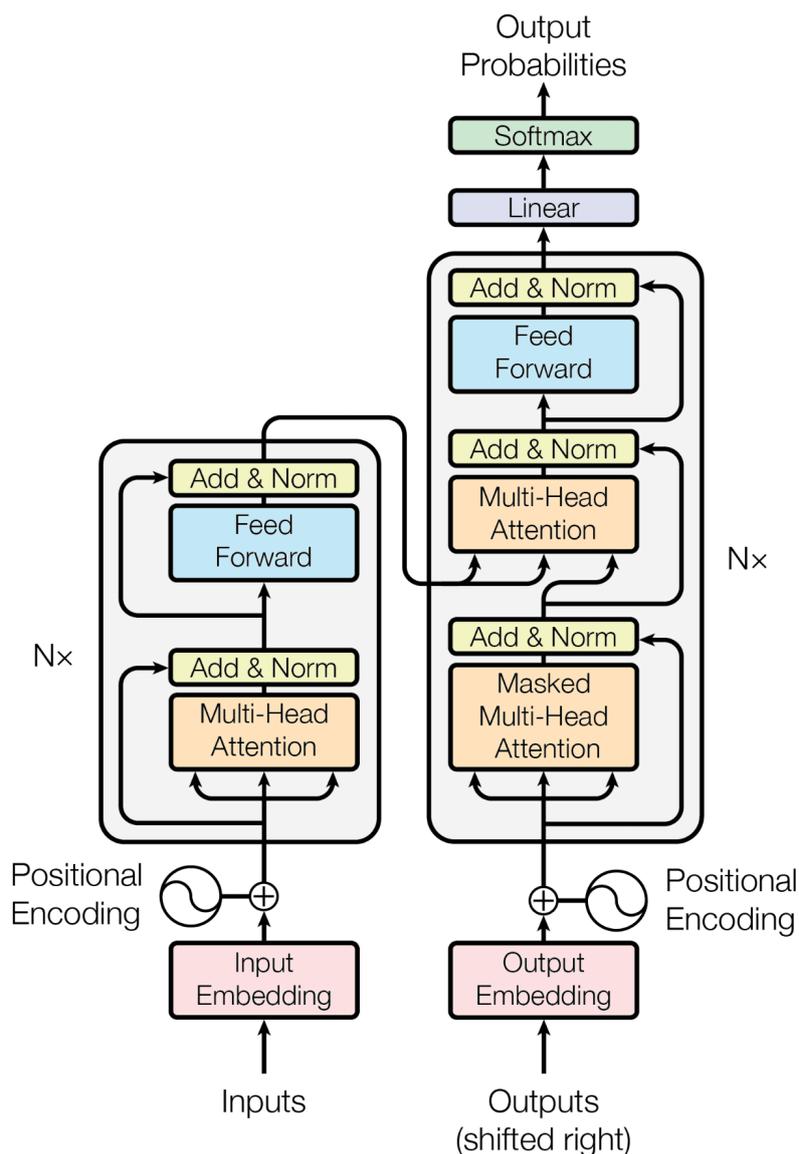


Figure I.7: Architecture du modèle de transformateur.[3]

la prédiction de la phrase suivante. La prédiction de la phrase suivante, qui permet au modèle de comprendre la relation entre les mots d'une phrase et entre les phrases.

Mise en œuvre : Les auteurs ont formé BERT sur un corpus massif de texte non étiqueté composé de 3,3 milliards de mots provenant de diverses sources, y compris des livres et des pages Web. Un objectif de pré-formation a été utilisé, qui consistait à masquer aléatoirement certains mots de la séquence d'entrée et à demander au modèle de prédire les mots masqués. Ensuite, la prochaine tâche de prédiction de phrase est introduite. Il s'agit d'alimenter le modèle en deux phrases et de lui demander de prédire si la deuxième phrase est la phrase suivante dans le texte original. Cette tâche permet au modèle d'en savoir plus sur les relations entre les phrases. BERT a été perfectionné sur une gamme de tâches NLP ciblées, y compris la réponse aux questions, l'analyse des sentiments et la

reconnaissance d'entités nommées, et a obtenu des résultats de pointe sur bon nombre de ces tâches.

Résultats : Les auteurs montrent que l'BERT les surpasse systématiquement sur une gamme de tâches NLP ciblées. Par exemple, sur le benchmark GLUE (General Language Understanding Evaluation), qui mesure les performances sur une gamme de tâches NLP, BERT a obtenu les meilleurs résultats sur les neuf tâches. BERT a obtenu un nouveau meilleur score sur le benchmark SQuAD, qui mesure les performances en matière de réponse aux questions.

Les auteurs montrent que l'BERT les surpasse systématiquement sur une gamme de tâches NLP ciblées. Par exemple, sur le benchmark GLUE (General Language Understanding Evaluation), qui mesure les performances sur une gamme de tâches NLP, BERT a obtenu les meilleurs résultats sur les neuf tâches. BERT a obtenu un nouveau meilleur score sur le benchmark SQuAD, qui mesure les performances en matière de réponse aux questions.

I.8.3 Approche basée sur le lexique

Dans cette approche, les mots d'opinion sont divisés en deux catégories. Les mots d'opinion positifs sont utilisés pour exprimer certaines choses nécessaires, et les mots d'opinion négatifs sont utilisés pour décrire des choses inutiles.[59]

I.8.3.1 Approche manuelle

Les mots d'opinion sont collectés manuellement sur la base de la connaissance du domaine et de la compréhension du langage des individus.[57]

I.8.3.2 Approche basée sur le corpus

L'utilisation de l'approche basée sur le corpus seul pour identifier tous les mots d'opinion n'est cependant pas aussi efficace que l'approche basée sur le dictionnaire car il est difficile de préparer un énorme corpus pour couvrir tous les mots anglais. Mais elle peut aider à trouver des mots d'opinion spécifiques au domaine et au contexte en utilisant un corpus de domaine, ce qui constitue l'énorme avantage de cette méthode. L'approche basée sur le corpus est réalisée selon une approche statistique ou une approche sémantique.[59]

I.8.3.3 Approche basée sur les dictionnaires

L'une des techniques simples de cette approche repose sur l'amorçage à l'aide d'un petit ensemble de mots d'opinion et d'un dictionnaire en ligne, par exemple WordNet. La stratégie consiste à collecter manuellement un petit ensemble de mots d'opinion dont l'orientation est connue, puis à élargir cet ensemble en recherchant leurs synonymes et antonymes dans le WordNet. Les mots nouvellement trouvés sont ajoutés à la liste de semences. L'itération

suivante commence. Le processus itératif s'arrête lorsque plus aucun nouveau mot n'est trouvé. Une fois le processus terminé, une inspection manuelle peut être effectuée pour supprimer et/ou corriger les erreurs.[59]

I.8.4 Approche hybride

L'approche hybride de l'analyse des sentiments exploite à la fois des méthodes statistiques et des méthodes basées sur la connaissance pour la détection de la polarité. Elle bénéficie de la précision élevée de l'apprentissage automatique (méthodes statistiques) et de la stabilité de l'approche basée sur le lexique.[60]

I.8.5 Autres Approches

Ces approches sont des approches d'analyse des sentiments qu'il est impossible de classer par catégorie.

I.9 Quelques travaux dans la prédiction des émotions

Dans le tableau I.2 nous présentons quelques travaux dans la prédiction des émotions. Les méthodes et les algorithmes utilisés, ainsi que les bases de données employées sont présentées.

Auteur	Méthode utilisée	Algorithmes utilisés	Dataset	Entrée	Sortie
Wanli Xing et al [61]	emotion-related (mot clé)	svm, arbre de décision, naïve base	Forum	les messages codés	Activation positive, Désactivation positive, Activation négative, Désactivation négative
Baker et al. [62]	observation data and log files	J48, JRip, Naive Bayes, K*	ASSISTments data	les demandes d'indices d'échafaudages, le contexte, le temps pris	Boredom, Frustration, Engaged, Concentration, Confusion, Off-Task
Sannyuya Liu, et al. [63]	manual coding	BERT-CNN	Forum de discussion en MOOC de 8867	Posts and achievement data	Emotional engagement (Positive-Confusion-Negative)
Imane Bouacida [2]	lexical-based sentiment classification	Latent Dirichlet Allocation (LDA), Latent Semantic Indexing (LSI)	Mark mon professeur (Hungary website)	Sac de mots $R_n * m$ matrix	Negative, Positive
Bowen Liu et al[64].	emotion related	K-Nearest, Logistic Regression, Naive Bayes, Decision Tree, Random Forest, SVM	MOOC forum	Posts in forum	Activation positive, Désactivation positive, Activation négative, Désactivation négative

Table I.2: Quelques travaux connexes

I.10 Conclusion

L'engagement des apprenants dans l'apprentissage en ligne est une préoccupation importante dans diverses études de recherche. Il joue un rôle crucial dans le processus d'apprentissage, car il a un impact significatif sur les résultats scolaires et la motivation des étudiants. Dans ce premier chapitre, nous avons présentés les concepts de bases liés au

sujet traité dans ce travail.

Le chapitre suivant, fait l'objet de la conception de notre systèmes.

II.1 Introduction

Le deuxième chapitre de ce travail est consacré à la conception de notre système. Nous décrivons dans ce chapitre une approche automatique pour la détection des émotions des apprenants. Le processus suivie par l'approche proposée est décrit en détail dans ce qui suit.

II.2 Objectif

Dans ce travail, on s'intéresse à l'exploration des avis pour pouvoir détecter automatiquement les émotions des apprenants. Pour cela, nous proposons une approche exploitant les techniques du Machine Learning pour détecter l'engagement émotionnel des apprenants dans les environnements d'apprentissage en ligne. Les avis des apprenants seront analysés pour pouvoir extraire les caractéristiques distinctives qui nous permettent de détecter les émotions des apprenants.

II.3 Description de l'approche proposée

Dans notre travail nous proposons une approche pour la détection automatique de l'engagement émotionnel à partir de l'exploration des avis des apprenants pour pouvoir détecter leurs émotions. L'approche proposée passe par quatre étapes, à savoir: le prétraitement, l'extract-ion de caractéristiques, l'apprentissage et la décision. Pour la décision, nous avons utilisé trois classifieurs issus des techniques de Machine learning: le naïve bayésienne (NB) et les Machines à Vecteurs de Support (Support Vector Machines: SVM), et le Bidirectional Encoder Representations from Transformers (BERT) qui est une technique du machine learning reposant sur le modèle de réseaux de neurones (Deep Learning). La figure II.1

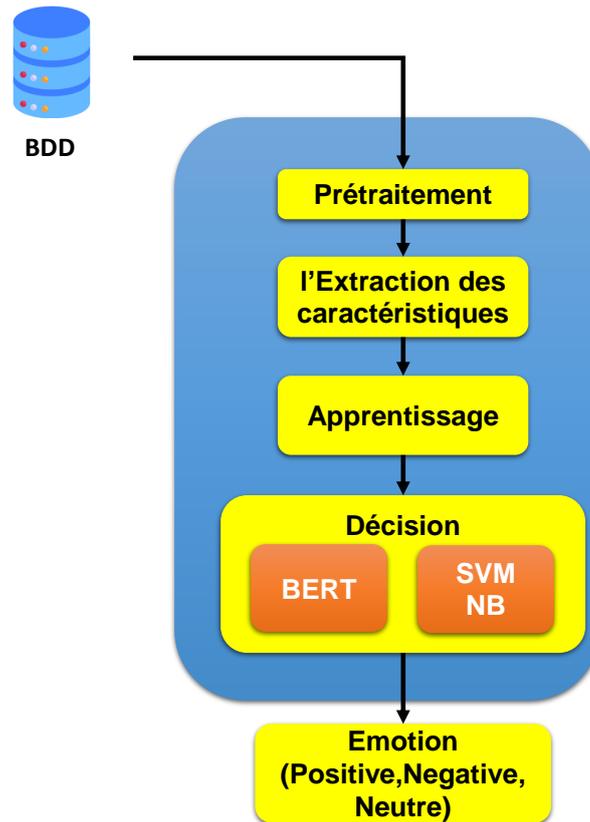


Figure II.1: L'approche proposée.

illustre le processus suivie l'approche proposée

II.4 Prétraitement

Le prétraitement dans l'apprentissage automatique fait référence aux mesures prises pour transformer les données brutes dans un format qui peut être utilisé par un modèle d'apprentissage automatique. Il s'agit généralement de nettoyer les données pour supprimer toute incohérence ou erreur. La figure II.2 illustre les étapes à suivre dans le prétraitement des données dans le cas du texte. À la fin du prétraitement nous écrivons le texte nettoyé dans la base de données.

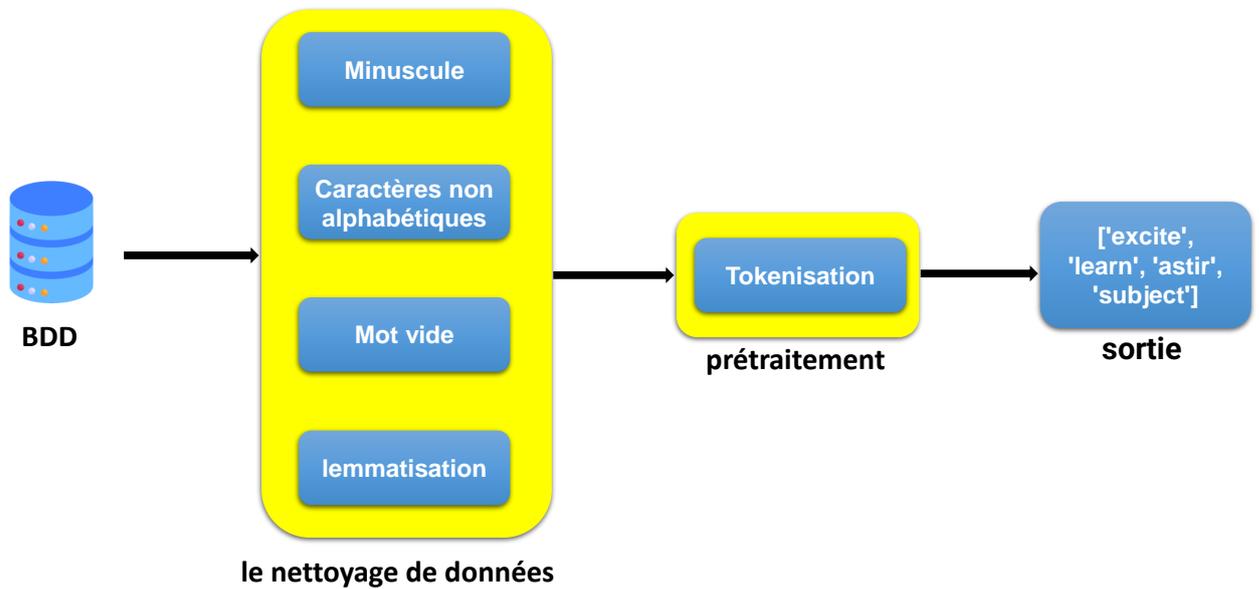


Figure II.2: Le prétraitement du Text (preprocessing)

II.4.1 Minuscule (Lowercase)

Nous faisons référence aux minuscules comme étant le processus de conversion de tous les caractères majuscules d'un texte donné en minuscules, cela peut être fait en utilisant la fonction `lower()` qui prend une chaîne en entrée et renvoie une nouvelle chaîne avec tous les caractères convertis en minuscules comme on peut le voir sur la figure II.3

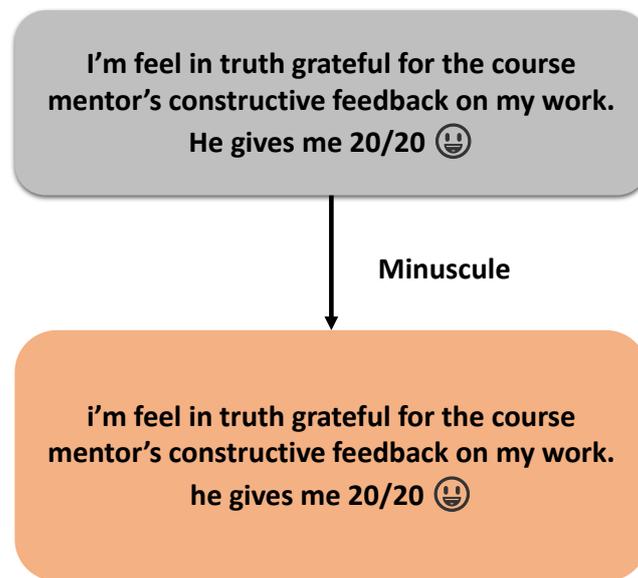


Figure II.3: miniscule

II.4.2 Caractères non alphabétiques(Non-alphabetic characters)

Nous utilisons le terme caractères non alphabétiques pour désigner tous les caractères d'un texte donné qui ne sont pas des lettres de l'alphabet. Il peut s'agir de la ponctuation, des chiffres, des espaces, des emojis, des nombres et d'autres caractères spéciaux comme on peut le voir sur la figure II.4

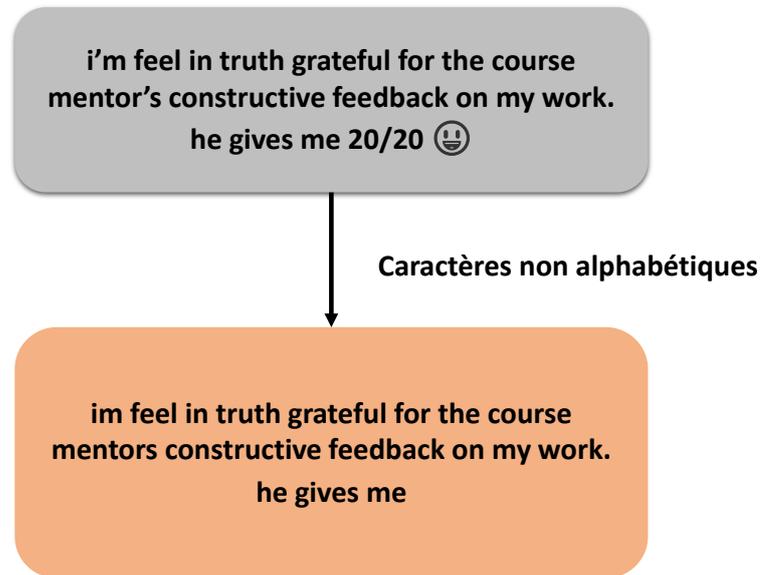


Figure II.4: Non-alphabetic

II.4.3 Mot vide (Stop words)

Nous supprimons souvent les mots vides, qui sont des mots courants, lors du traitement de texte en langage naturel, car ils sont considérés comme ayant peu de valeur sémantique et peu susceptibles de contribuer au sens général d'un texte. Parmi les exemples de mots vides, nous citons "the", "and", "a", "in", and "is". La figure II.5 montre un exemple de cette étape.

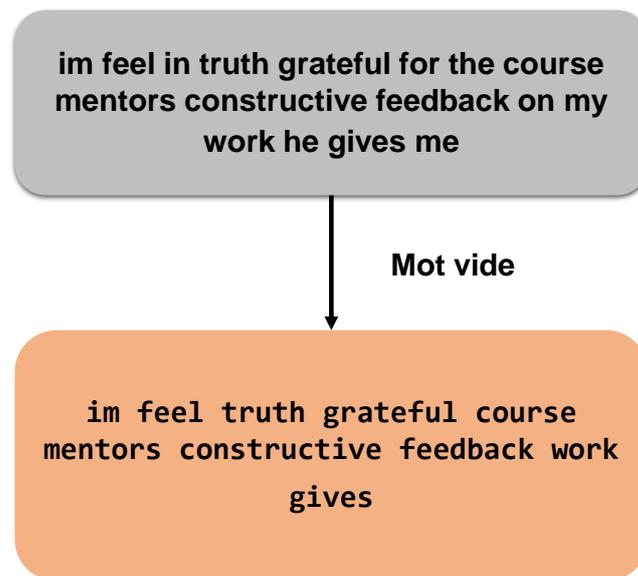


Figure II.5: Mot vide

II.4.4 lemmatisation(Lemmatization)

Nous avons réalisé la lemmatisation, qui est le processus de réduction d'un mot à sa forme de base ou racine, connue sous le nom de lemme. Cette opération est souvent réalisée dans le cadre de tâches de traitement du langage naturel afin de réduire la dimensionnalité de l'espace des caractéristiques et de normaliser le texte comme on peut le voir sur la figure II.6

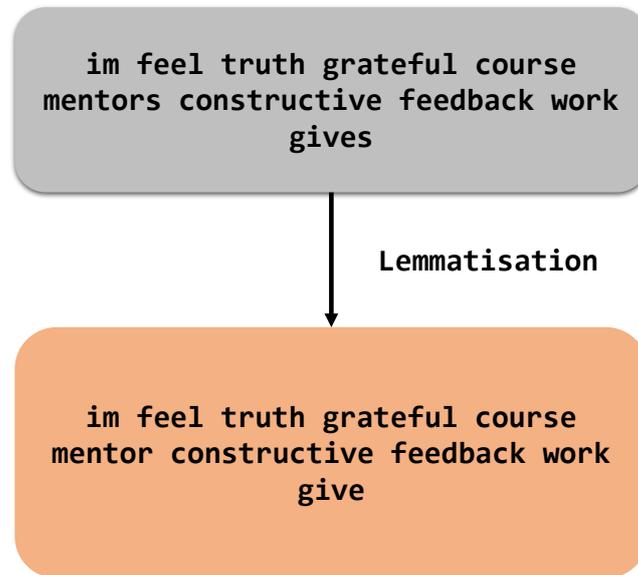


Figure II.6: lemmatisation

II.4.5 Tokénisation(Tokenization)

Nous effectuons la tokenisation en divisant un texte en unités plus petites, appelées tokens (La tokenisation est le processus qui consiste à diviser un texte en unités plus petites, appelées tokens). Ces jetons peuvent être des mots, des phrases ou même des caractères individuels, en fonction des besoins spécifiques de la tâche de traitement du langage naturel. La tokenisation est une étape importante qui permet de transformer un texte non structuré en un format structuré qui peut être plus facilement analysé et traité comme on peut le voir sur la figure II.7

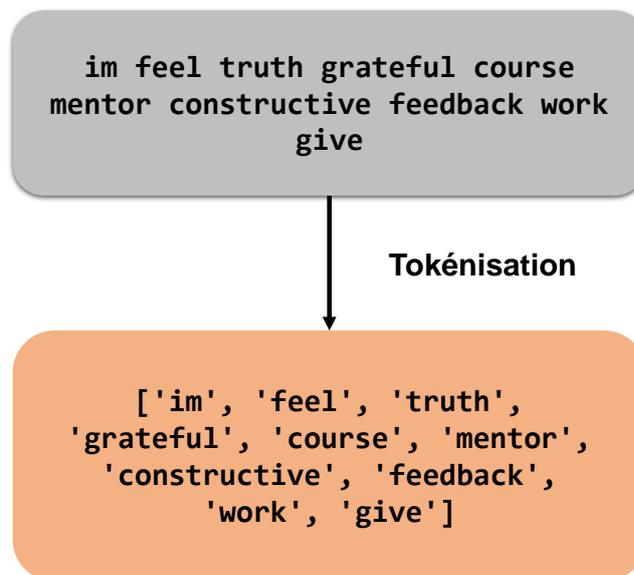


Figure II.7: Tokénisation

Les étapes de prétraitement présentées précédemment ont été faites pour les deux premiers classifieurs (NB et SVM). Dans ce qui suit, nous expliquons le prétraitement qui a été fait pour le modèle de BERT.

Dans le contexte de BERT, le prétraitement se réfère aux étapes prises pour préparer les données d'entrée avant de les introduire dans le modèle pour la classification.

Les étapes de prétraitement pour BERT impliquent généralement la tokenisation, l'ajout de tokens spéciaux et la génération d'embeddings d'entrée. Voici une explication étape par étape du processus de prétraitement de BERT :

- **Tokénisation** : Le texte d'entrée est divisé en tokens individuels. BERT utilise la tokenisation WordPiece, qui décompose les mots en sous-mots et les représente à l'aide d'un vocabulaire d'unités de sous-mots. Par exemple, le mot "playing" peut être transformé en "play" et "##ing". Les jetons sont généralement en minuscules.
- **Jetons spéciaux** : Des jetons spéciaux sont ajoutés pour marquer le début et la fin du texte d'entrée, ainsi que pour séparer les phrases ou les segments dans le texte. BERT utilise généralement le jeton [CLS] pour représenter le début de la séquence et le jeton [SEP] pour séparer les phrases ou les segments.
- **Remplissage et troncature** : Pour garantir la cohérence de la longueur d'entrée,

les séquences sont complétées ou tronquées à une longueur fixe. BERT utilise une longueur de séquence maximale, et si une séquence est plus longue que la longueur spécifiée, elle est tronquée. Si elle est plus courte, des jetons de remplissage, généralement [PAD], sont ajoutés pour lui donner la même longueur.

- **Encastremements d'entrée** : BERT nécessite trois types d'enregistrements d'entrée : enregistrements de jetons, enregistrements de segments et enregistrements de positions.
 - **Emboîtements de jetons** : Chaque jeton se voit attribuer un identifiant unique basé sur le vocabulaire de BERT. Ces ID de jeton sont utilisés pour rechercher les enchâssements de jeton correspondants dans le modèle BERT pré-entraîné. Les encastremements de jetons capturent la signification de chaque jeton.
 - **Segment Embeddings** : BERT prend en charge la classification de textes comportant plusieurs segments ou phrases. Pour faire la distinction entre ces segments, des enregistrements de segments sont ajoutés. Chaque jeton se voit attribuer un identifiant de segment indiquant son appartenance à une phrase ou à un segment.
 - **Encastremements de position** : Les encastremements de position codent la position de chaque token dans la séquence. Ils sont ajoutés pour fournir des informations de position au modèle.

Une fois que les données d'entrée ont subi ces étapes de prétraitement, elles peuvent être introduites dans le modèle BERT pour la classification. La figure II.8 illustre le prétraitement du BERT.

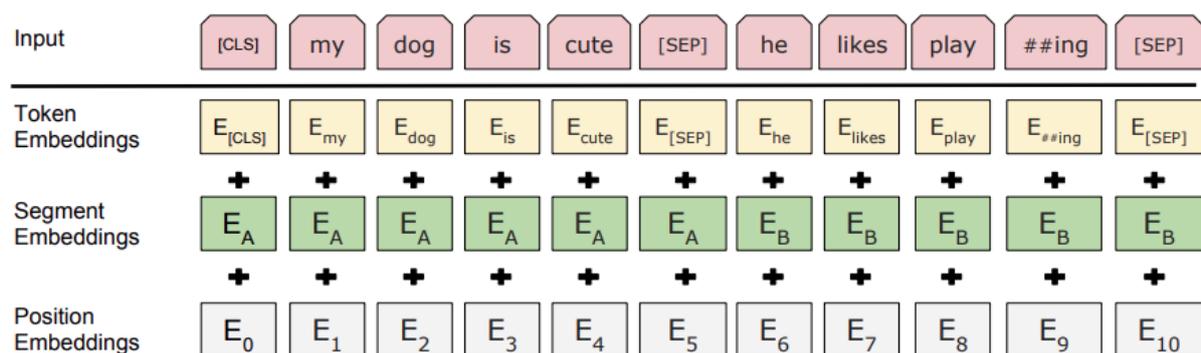


Figure II.8: Le prétraitement du BERT.

II.5 Extraction de caractéristiques: Bag of Words (BOW)

L'extraction de caractéristiques est une étape cruciale dans les systèmes de classification ou de reconnaissance. Pour représenter un document texte d'entrée sous une forme numérique,

des caractéristiques distinctives doivent être extraites. Dans cette section, nous présentons les caractéristiques que nous avons utilisé dans notre système de détection automatique de l'engagement émotionnel. Dans notre système, nous utilisons le sac de mots (en anglais Bag of Words (BoW)). Le modèle de sac de mots est utilisé pour analyser un texte en le convertissant en un sac de mots ou un vecteur de longueur fixe. Le vecteur représente les occurrences (fréquence) des mots dans le document.

Dans notre cas, nous avons utilisé deux sacs d'adjectifs positifs et négatifs, chaque sac contient les 20 premiers adjectifs du WordNet Voir la figure II.10. WordNet est une base de données lexicale qui répertorie, classifie et met en relation de diverses manières le contenu sémantique et lexicale de la langue anglaise. L'idée sous-jacente de WordNet est de créer un "lexique de sens" qui combine les fonctions d'un dictionnaire et d'un thésaurus. Au lieu d'organiser les informations lexicales en fonction des formes de mots, elles sont classées selon les significations des mots, en harmonie avec les représentations humaines du sens et leur traitement dans le cerveau[2].

Après l'extraction des mots (positifs ou négatifs) des messages ou commentaires de notre base de données, nous ne nous limitons pas à compter les adjectifs figurant dans ces commentaires, mais le vecteur est enrichi avec les synonymes de ces adjectifs à l'aide du WordNet.

Par la suite, et pour augmenter la pertinence des mots clés utilisés, notre vecteur de mots est convertit en un vecteur numérique en utilisant la méthode de pondération TF-IDF (Term Frequency-Inverse Document Frequency). La formule de la méthode TF-IDF est la suivante:

La figure II.9 illustre les étapes pour créer le vecteur pour chaque commentaire. Dans l'exemple suivant, nous montrons un exemple d'un extrait du vecteur de caractéristiques.

II.6 Apprentissage

Le but de cette phase est d'estimer un modèle à partir des caractéristiques retenues. A la fin de l'apprentissage, des fichiers modèles sont générés.

II.7 Décision

La décision est la dernière étape de notre processus de détection d'émotions. Trois classes de sorties sont définies, à savoir: Positive, Négative, et Neutre. le tableau suivant représente le schéma d'encodage des trois émotions. Pour la décision, nous avons utilisé trois techniques du Machine Learning à savoir: classification Naive Bayésienne (NB), les Machines à Vecteurs de Supports (SVM) et le Bidirectional Encoder Representations from Transformers (BERT).

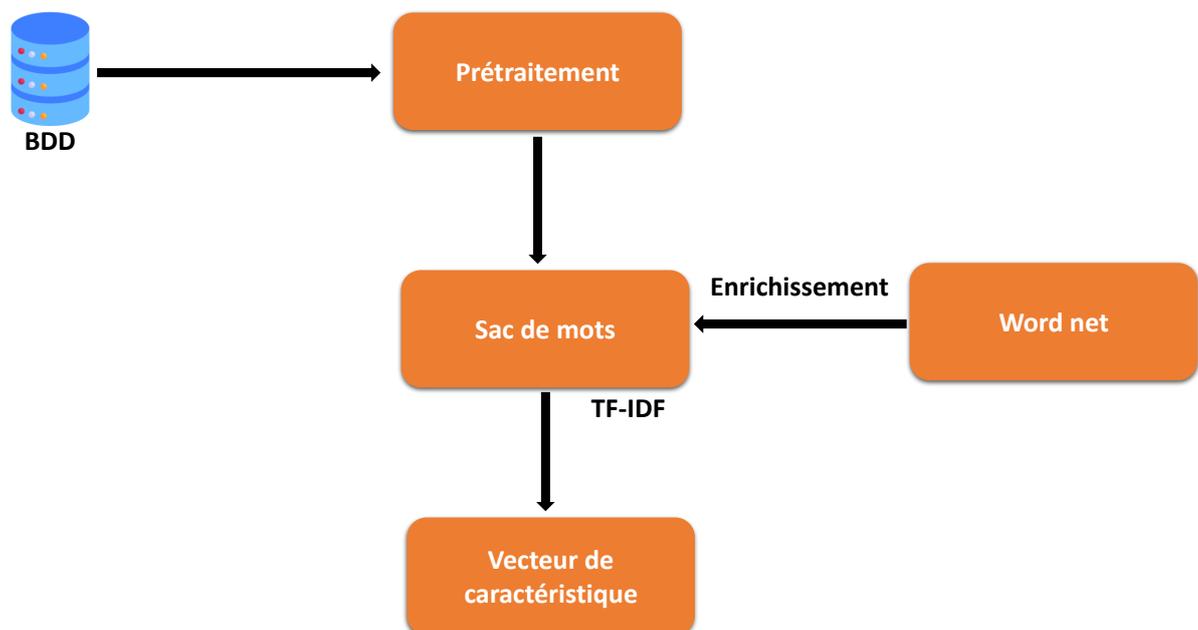


Figure II.9: Construction du vecteur de caractéristique.



Figure II.10: Vecteur positive et negative adjective

II.7.1 La classification Naive bayésienne (NB)

Les classifieurs de Naive Bayes (NB) sont une famille de classificateurs basés sur le célèbre théorème de probabilité de Bayes, connus pour créer des modèles simples et puissants, particulièrement dans les domaines de la classification des documents et de la prédiction des maladies. La classification textuelle de NB est le plus souvent utilisée pour catégoriser des textes, car elle est rapide et facile à mettre en oeuvre [65]. Ce type de classification simple permet à un modèle de machine learning d'apprendre rapidement. Il n'est pas nécessaire de fournir un gros volume de données lors de la phase d'apprentissage.

Engagement Émotionnelle	Indicateur	Definition	Exemple
	Positive	L'apprenant exprime des émotions positives telles que fierté, surprise, souhait, etc.	<ul style="list-style-type: none"> - The lecture he held, and he was my intern, was very helpful all the time. - Useful subject and interesting presentation. - A very good teacher and always a good head for everyone.
	Negative	L'apprenant exprime des émotions négatives telles que débordement, frustration, etc.	<ul style="list-style-type: none"> - Boring presentation mode. - I did not really understand what this man would want. - His lectures were not interesting.
	Neutre	L'apprenant n'exprime aucune émotion.	<ul style="list-style-type: none"> - Caffeine-based lifestyle. - Moderated. - Hello.

Table II.1: Le schéma d'encodage de l'engagement émotionnel.

II.7.2 Les Machines à Vecteurs de Support

Contrairement aux classifieurs classiques qui se base sur la minimisation du risque empirique, et qui ne garantit pas la convergence vers une solution optimale, les SVM opèrent sur le principe de la minimisation du risque structurel. Ce risque constitue une borne supérieure du risque réel ou l'erreur d'apprentissage. Le pouvoir des SVM réside dans la construction d'un hyperplan optimal de séparation entre les données d'apprentissage [66]. L'algorithme de construction du SVM cherche un hyperplan de séparation entre les classes à discriminer, dont la distance minimale aux exemples d'apprentissage est maximale (voir la figure). La classe des fonctions linéaires n'est pas suffisamment discriminante, et choisir des frontières de décision linéaires semble être un facteur limitant. Cependant, de tels modèles peuvent être considérablement enrichis en projetant (généralement non-linéairement) les données dans un espace augmenté (feature space), éventuellement de plus grande dimension que l'espace des entrées, afin de rendre linéairement séparable le jeu de données.

II.7.3 Bidirectional Encoder Representations from Transformers (BERT)

Le modèle de Bidirectional Encoder Representations from Transformers (BERT) de Devlin et al [58] utilise les encodeurs dans un transformateur comme une sous-structure pour pré-entraîner des modèles pour des tâches de NLP telles que: Question Answering (QA), Text Summarization (TS), etc. L'exécution de BERT pour ces tâches se fait en deux phases, à savoir le pré-entraînement pour la compréhension du langage et le réglage fin pour une

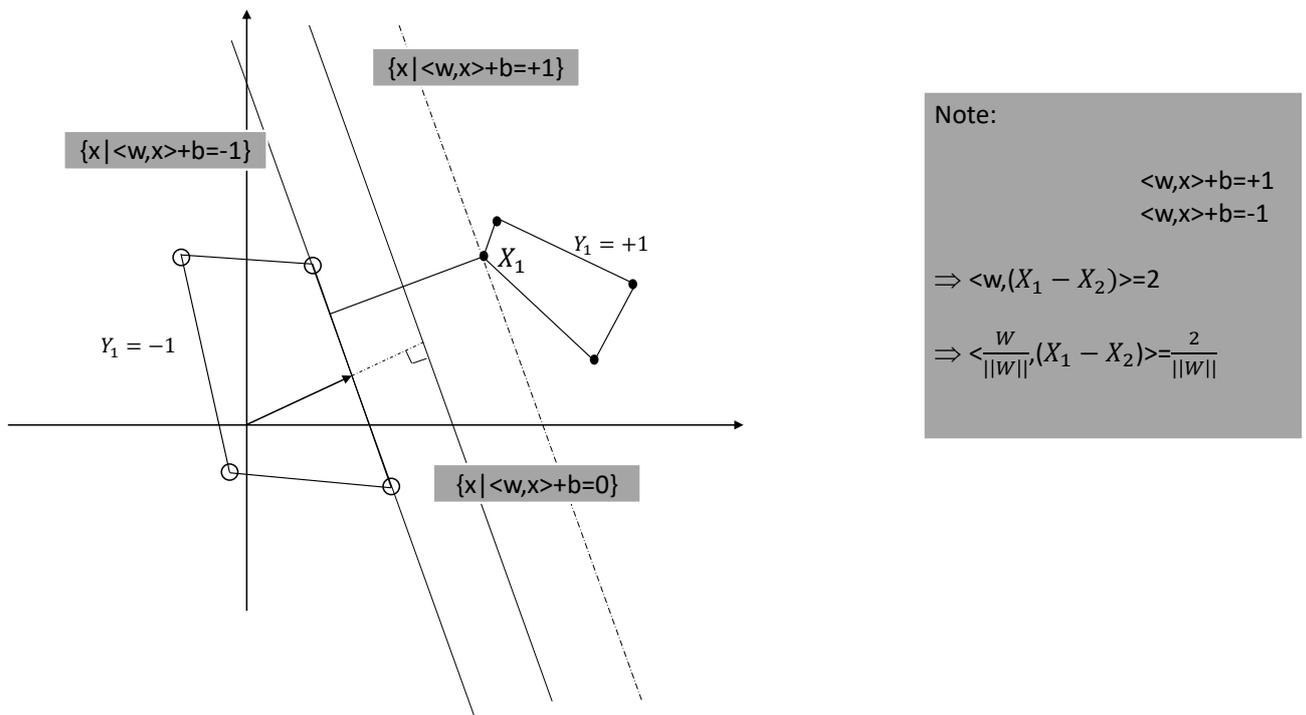


Figure II.11: Hyperplan de séparation

tâche spécifique.

BERT peut comprendre le langage en s'entraînant sur les mécanismes de modélisation du langage masqué (MLM) et de prédiction de la phrase suivante (PPS). BERT suppose un blinder avec MLM pour apprendre les contextes bidirectionnels dans les phrases. Ainsi, il prend en entrée des phrases aléatoires, masque certains mots dans les phrases et reconstruit les mots masqués à partir des textes environnants en sortie. Sa capacité vient du fait qu'il permet de saisir deux phrases à la fois et à déterminer si la seconde phrase vient après la première. si la deuxième phrase vient après la première lui permet d'atteindre le PPS. Cette capacité aide le de maintenir des relations à longue distance entre les textes. BERT a été entraîné sur des textes de 16 Go provenant des ensembles de données BooksCorpus (Zhu et al [67].) et de la Wikipédia anglaise.

Après le pré-entraînement, le modèle est ensuite entraîné sur une tâche de NLP en effectuant un entraînement supervisé sur un ensemble de données et en remplaçant la sortie entièrement connectée du BERT par un nouvel ensemble de couches de sortie. Le modèle BERT s'entraîne plus rapidement car les autres paramètres du modèle ne sont réglés que de manière ponctuelle, à l'exception des paramètres de sortie appris à partir de zéro (Fig. II.12).

Le BERT a deux modèles, à savoir le modèle BERT-base et le modèle BERT-Large II.13 . Le modèle BERT-base comprend des blocs de codage à transformateur à 12 couches,

chaque bloc contenant 12 couches d'auto-attention et 768 couches cachées et produisant \approx 110 millions de paramètres au total. D'autre part, le modèle BERT-large est composé de 24 blocs d'encodage à transformateur en couches, Chaque bloc contient 24 couches d'auto-attention de tête et produit un total de \approx 340 millions de paramètres. Les performances de l'BERT dépendent du type de modèle c'est-à-dire que BERT-large peut atteindre des précisions plus élevées que BERT-base. Cependant, cette amélioration de la précision en utilisant le BERT-large se fait au prix de ressources plus importantes. L'architecture générale du modèle BERT est illustrée à la figure II.14 .

Les avantages de l'BERT comprennent sa capacité à gérer l'extraction d'informations contextuelles grâce à sa capacité bidirectionnelle ; elle s'entraîne plus rapidement et a été utilisée dans un large éventail d'applications de modélisation linguistique. Cependant, les inconvénients suivants persistent pour le BERT : il est limité aux classifications monolingues, la longueur des phrases d'entrée le limite également, il souffre de l'inférence pragmatique, et l'utilisation du BERT-large peut avoir tendance à être coûteuse en termes de calcul.

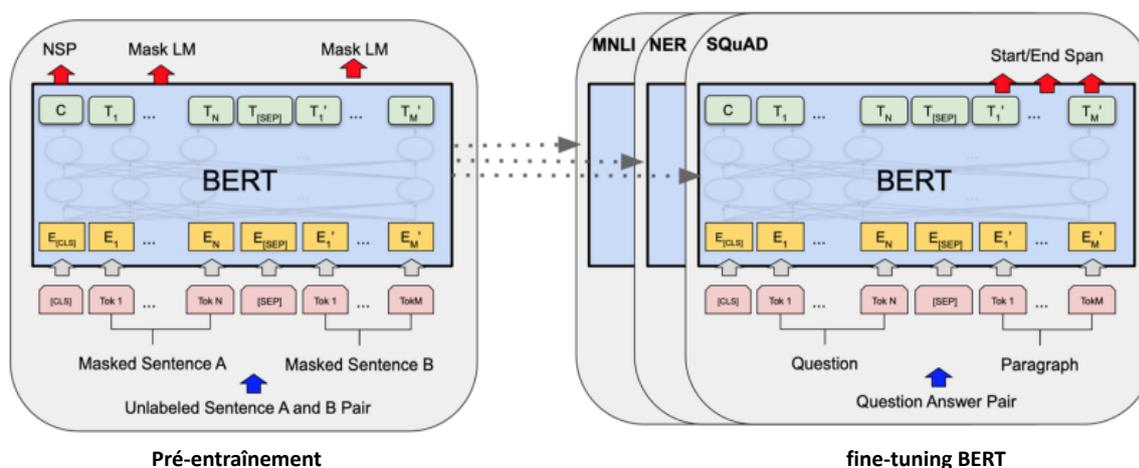


Figure II.12: Pré-entraînement et fine-tuning BERT.[4]

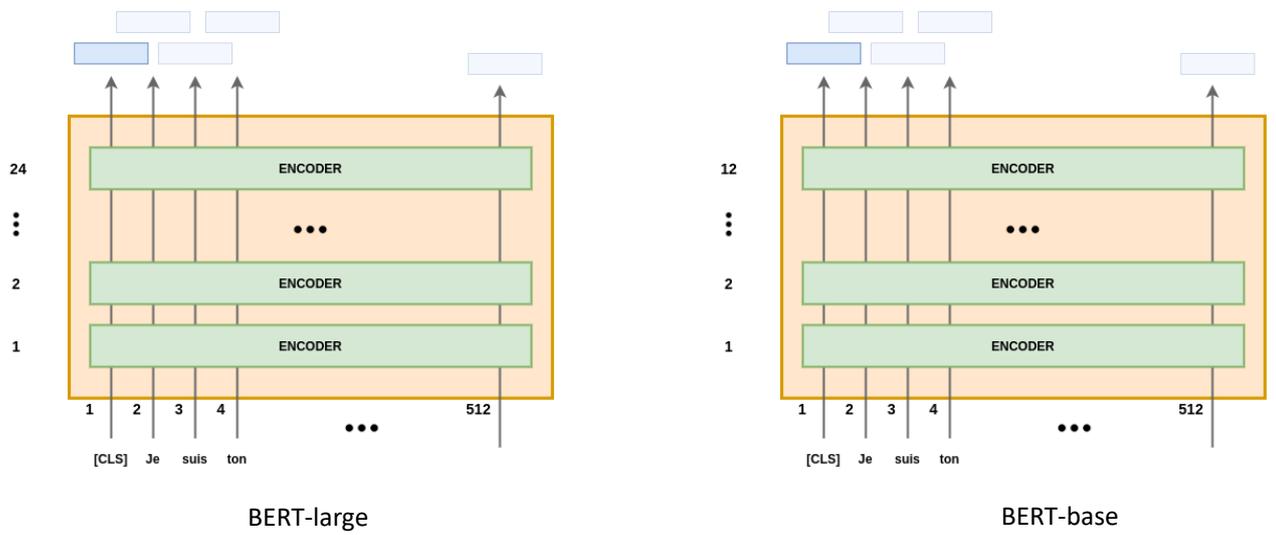


Figure II.13: BERT-large et bert-base.[5]

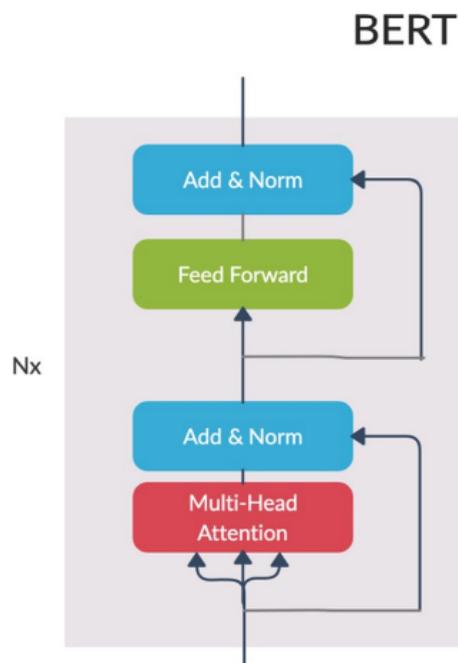


Figure II.14: Architecture du modèle BERT.[4]

II.8 Conclusion

Le deuxième chapitre de ce mémoire a été consacré à la conception de notre système. Nous avons présenté en détail les étapes de notre processus de détection automatique des émotions des apprenants à partir des leurs avis.

Dans le chapitre suivant, nous verrons les résultats de chaque approche et nous ferons la comparaison entre les approches et nous saurons quelle est la meilleure approche pour faire de la détection d'émotions.

III.1 Introduction

Après avoir expliqué l’approche proposée en détail dans le chapitre précédent, nous allons présenter, dans ce chapitre, l’implémentation de notre système, ainsi que l’environnement de développement. La deuxième partie de ce chapitre est consacrée au test du modèle et aux résultats obtenus, et une comparaison entre les deux approches proposées.

III.2 Environnement

III.2.1 Google Colaboratory

Colaboratory, souvent abrégé en ”Colab”, est un produit de Google Research. Colab permet à quiconque d’écrire et d’exécuter le code Python de son choix via un navigateur. Il s’agit d’un environnement particulièrement adapté à l’apprentissage automatique, à l’analyse de données et à l’éducation. En termes plus techniques, Colab est un service de bloc-notes Jupyter hébergé qui ne nécessite aucune configuration et offre un accès gratuit aux ressources informatiques, y compris les GPU. La figure III.1 présente l’interface de Colab. En termes plus techniques, Colab vous permet de :

- Écrire et exécuter du code Python : Google Colab fournit une interface de bloc-notes dans laquelle vous pouvez écrire et exécuter du code Python. Vous pouvez écrire et exécuter des cellules de code individuellement ou exécuter l’ensemble du carnet.
- Collaborer et partager : Vous pouvez facilement partager vos carnets Colab avec d’autres personnes, ce qui leur permet de voir et de modifier votre code.
- Analyse et visualisation des données : Google Colab fournit plusieurs bibliothèques

et outils pour l'analyse et la visualisation des données, tels que Pandas, NumPy, Matplotlib et Seaborn.

- Apprentissage automatique et apprentissage profond : Colab offre une prise en charge intégrée des bibliothèques d'apprentissage automatique et d'apprentissage profond les plus répandues, telles que TensorFlow et PyTorch.
- Accès à des ressources supplémentaires : Google Colab vous permet d'installer des bibliothèques Python et des dépendances supplémentaires en fonction de vos besoins.

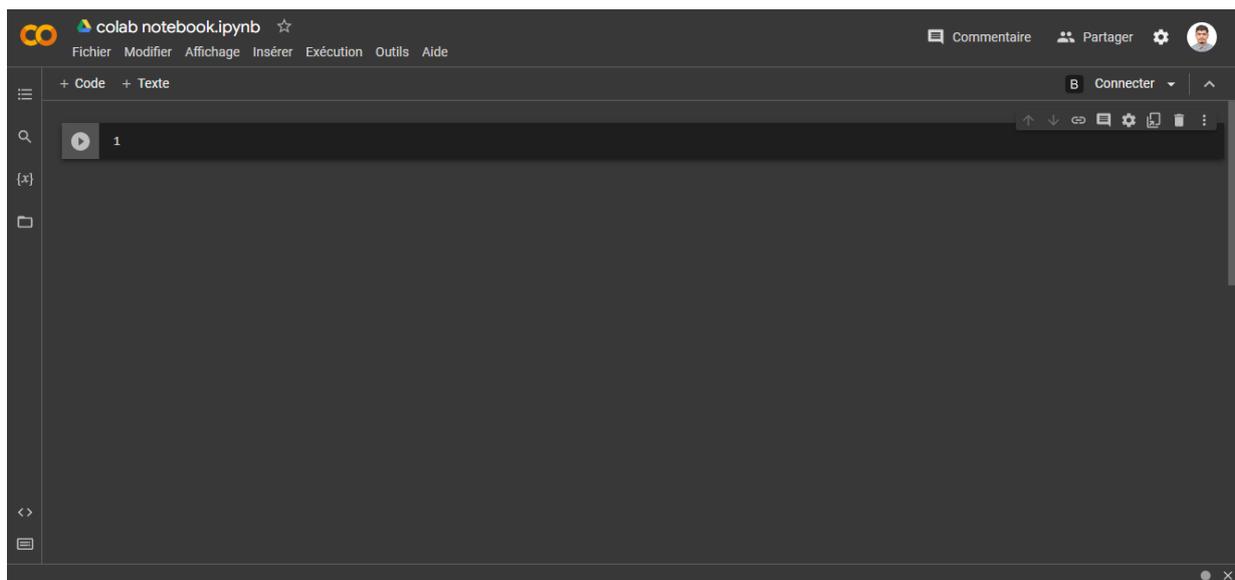


Figure III.1: Interface de Colab.

III.2.2 Apprentissage sur GPU

Pour configurer le GPU (processeur graphique), il suffit de cliquer sur Exécution → Modifier le type d'exécution et sélectionner GPU comme accélérateur matériel. La figure III.2 présente l'activation de GPU.

III.3 Langage de programmation et bibliothèque utilisée

III.3.1 Python

Python est un langage de programmation populaire de haut niveau à usage général. Sa philosophie de conception met l'accent sur la lisibilité du code. Python prend en charge de nombreux paradigmes de programmation, y compris les styles orientés objet, impératifs et fonctionnels ou procéduraux. Il dispose d'un système de type dynamique et d'une gestion automatique de la mémoire, ainsi que d'une bibliothèque standard étendue et complète.

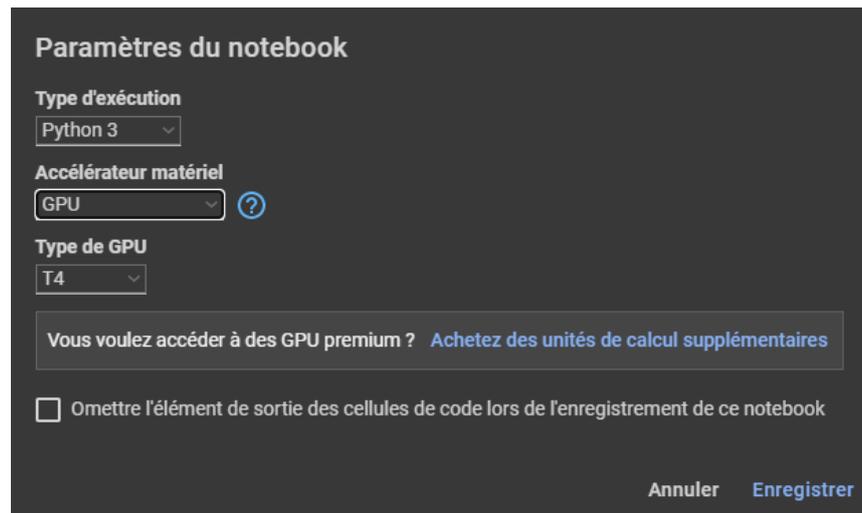


Figure III.2: Activation de GPU.

Dans l'ensemble, les langages professionnels sont le meilleur choix, mais en matière de programmation, Python arrive en deuxième position. Python possède également des bibliothèques tierces open source matures, telles que Numpy/Scipy pour les opérations numériques, Cython pour l'optimisation, Cython pour l'optimisation de bas niveau, IPython pour le travail interactif et matplotlib pour le dessin.[68]

III.3.2 NLTK

NLTK (Natural Language Toolkit) est l'une des bibliothèques Python les plus utilisées dans le domaine du traitement du langage naturel. NLTK est une bibliothèque Python qui permet de traiter rapidement et facilement des textes en langage naturel. La boîte à outils a été développée à l'Université de Pennsylvanie. Université de Pennsylvanie comme outil de recherche et d'enseignement pour le traitement du langage naturel. NLTK dispose d'un grand nombre de corpus intégrés, comprenant divers types de textes tels que des romans, des nouvelles, des textes de chat en réseau, des critiques de films, etc. Il comprend le Brown's Corpus, le Gutenberg's Corpus, le Inaugural Address Corpus, le Reuters Corpus [69]. En outre, NLTK fournit des interfaces faciles à utiliser pour plus de 50 corpus et ressources lexicales telles que WordNet, ainsi qu'une suite de bibliothèques de traitement de texte pour la classification, la tokenisation, le stemming, le tagging, l'analyse syntaxique et le raisonnement sémantique, des wrappers pour des bibliothèques NLP de qualité industrielle, et un forum de discussion actif. En collaboration avec la puissante bibliothèque standard de Python et d'autres bibliothèques tierces, il peut effectuer un traitement secondaire des résultats du traitement. Il constitue une base solide pour le traitement de textes complexes[69].

III.3.3 TensorFlow

TensorFlow est une bibliothèque logicielle flexible et extensible pour le calcul numérique à l'aide de graphes de flux de données. La bibliothèque et les outils associés permettent aux utilisateurs de programmer et de former efficacement des réseaux de neurones et d'autres modèles d'apprentissage automatique, et de les déployer en production. Les algorithmes de base de TensorFlow sont écrits en C++ hautement optimisé et en Compute Unified Device Architecture (CUDA), une plate-forme informatique parallèle et une API créées par NVIDIA. Il a des API dans plusieurs langues. L'API Python est la plus complète et la plus stable. Les autres langages officiellement pris en charge sont JavaScript, C++, Java, Go et Swift. Des packages tiers sont disponibles pour d'autres langages tels que C # et Ruby. Les exemples donnés dans cette revue sont basés sur Python, car c'est le langage le plus facile à lire et à utiliser[70].

III.3.4 Numpy

NumPy, abréviation de Numerical Python, est une bibliothèque composée d'objets de type tableau multidimensionnel et d'un ensemble de routines pour manipuler ces tableaux. NumPy permet d'effectuer des opérations mathématiques et logiques sur des tableaux[71].

III.4 La base de données (Data Set)

III.4.1 Description

Afin d'évaluer l'approche proposée, nous avons besoin d'une base de données. Dans notre cas, nous avons besoin d'un ensemble de données textuelles d'opinion qui peuvent être utilisées pour analyser l'apprentissage afin de détecter les émotions et les opinions des apprenants sur divers sujets d'intérêt et sur les cours présentés. Vu l'absence de Data Set, nous avons utilisé la base de données collectée à partir du site web Mark My Professor [2] . Cette base de données est constituée de 5200 avis d'apprenants concernant les cours présentés par leurs professeurs, 3372 avis positive, 982 avis negative et 846 avis neutre. La figure III.3 montre les dix premières lignes de la base de données .

III.4.2 Répartition de la base de données

Pour valider l'approche proposée, la base de données a été dévisée en 3 parties, à savoir: l'apprentissage, la validation, et le test. la répartition a été faite comme suit: 80% pour l'apprentissage , 10% pour la validation et 10% pour le test , le tableauIII.1 illustre la répartition de la base de données.

Paramètres du modèle	Valeurs
SVM	
C	[0.001, 0.01, 0.1, 1, 10, 100]
Gamma	[0.1, 0.01, 0.001, 0.0001]
Kernel	['poly', 'rbf', 'linear']
NB	
Param Grid	[0.1, 1.0, 10.0, 100.0, 1000.0, 10000.0]
Text Siz	0.2
Random State	32
BERT	
BATCH SIZE	16
EPOCHS	10
MAX SEQ LEN	512

Table III.2: Les valeurs des paramètres du modèle

mais il peut être augmenté ou diminué en fonction de la convergence du modèle. Dans notre model l'EPOCHS a été fixé 1.

- **MAX SEQ LEN** : la longueur maximale de la séquence fait référence au nombre maximal de jetons autorisés dans une seule séquence d'entrée. BERT nécessite des séquences d'entrée de longueur fixe pour un traitement efficace. La longueur de séquence maximale par défaut de BERT est souvent fixée à 512 jetons. Dans notre model **MAX SEQ LEN** a été 300.

Les valeurs par défaut des trois paramètres cités précédemment sont fixés à:

- **BATCH SIZE** = 8
- **EPOCHS** = 1
- **MAX SEQ LEN** = 300

Avec ces valeurs de paramètres, le temps d'apprentissage du modèle est de 40 minutes, la précision est à 10, la perte est à 10 et la perte de validation est à 10.

Vu ces résultats, les valeurs des paramètres ont été fixé d'une manière expérimentale comme suit:

- **BATCH SIZE** = 16
- **EPOCHS** = 10
- **MAX SEQ LEN** = 512

Avec ces paramètres, le temps d'apprentissage du modèle dans chaque epoch est de 9 minutes. Le graphe III.4 présente les variations de la précision (Accuracy) en fonction des Epochs. Le graphe III.5 présente Loss et Validation Loss en fonction des Epochs.

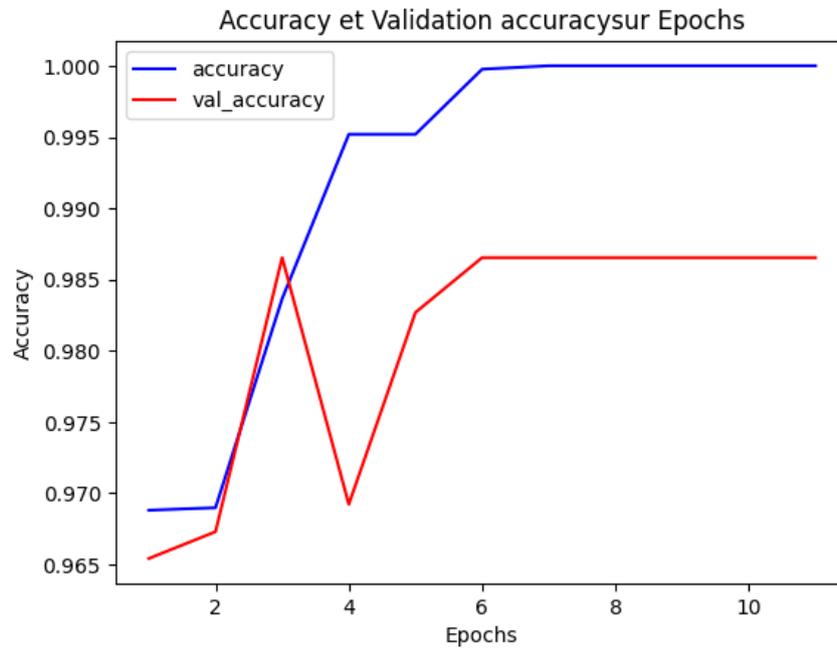


Figure III.4: Accuracy et Validation Accuracy sur Epoch.

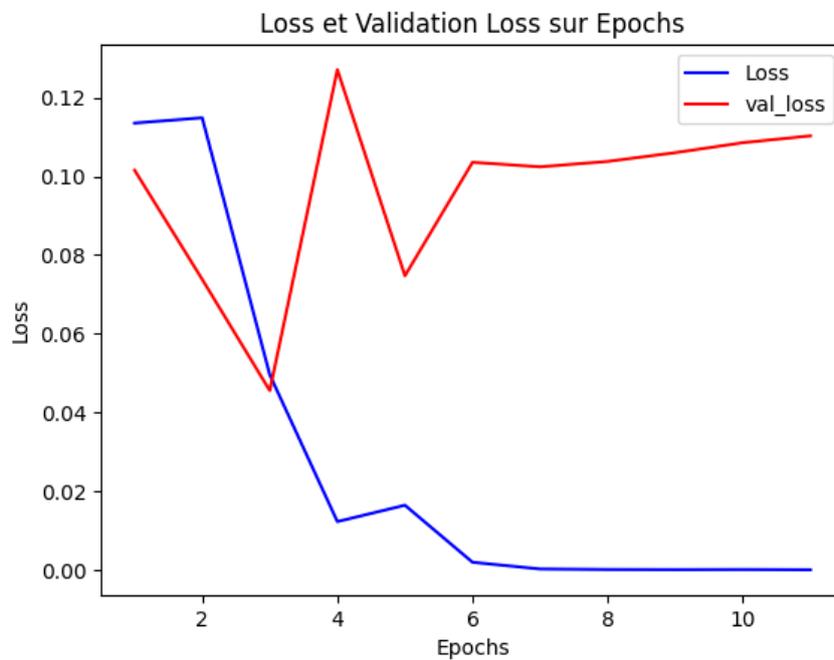


Figure III.5: Loss et Validation Loss sur Epochs.

III.6 Les résultats obtenus

Dans ce qui suit, nous présentons les résultats que nous obtenus en utilisons les trois classifieurs choisies. Pour évaluer ces classifieurs, il est nécessaire de calculer les paramètres suivants:

- **Vrai positif (VP):** Les cas où on prédit OUI et la sortie réelle est également OUI.
- **Vrai négatif (VN):** Les cas où on prédit NON et la sortie réelle est également NON.
- **Faux positif (FP):** les cas où on prédit OUI et la sortie réelle est NON.
- **Faux négatif (FN):** Les cas où on prédit NON et la sortie réelle est OUI.

En fonction de ces paramètres, les métriques de performances que nous avons utilisé sont les suivants:

- **La précision(Accuracy):** est l'habilité du classifieur de trouver uniquement les instances pertinente. Elle peut être calculée comme suit:

$$\text{Recall} = \frac{\text{VP}}{\text{VP} + \text{FP}}$$

- **Le rappel (Recall):** Le rappel est l'habilité du classifieur de classifier les instances correctement. Il peut être calculer via la formule suivante:

$$\text{Recall} = \frac{\text{VP}}{\text{VP} + \text{FN}}$$

- **F-Measure:** La f-mesure est définie comme une moyenne harmonique de la précision et du rappel. Elle peut être définie comme suit:

$$\text{F-measure} = 2 \times \frac{\text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$$

Les tableaux III.3 représente les résultats obtenus en utilisant le classifieur SVM.

Algorithme	Accuracy	Precision	Recall	F1-Score
SVM	0.68	0.23	0.33	0.26

Table III.3: Résultats de la classification SVM

Le tableau III.4 représente les résultats obtenus en utilisant le classifieur NB.

Le tableau III.5 présente les résultats obtenus en utilisant le classifieur BERT.

Algorithme	Accuracy	Precision	Recall	F1-Score
NB	0.74	0.76	0.74	0.67

Table III.4: Résultats de la classification NB

Algorithme	Accuracy	Precision	Recall	F1-Score
BERT	1	0.99	0.99	0.99

Table III.5: Résultats de la classification BERT.

III.7 Comparaison entre l'approche de Machine Learning et l'approche de Deep Learning

Comme nous pouvons le voir sur la figure III.6, la comparaison entre les techniques de classification SVM, NB et BERT en fonction de: accuracy, precision, recall et f1-score montre que le modèle de BERT donne les meilleurs résultats. Donc, nous pouvons conclure que BERT est la meilleur classifieur parmi trois utilisé.

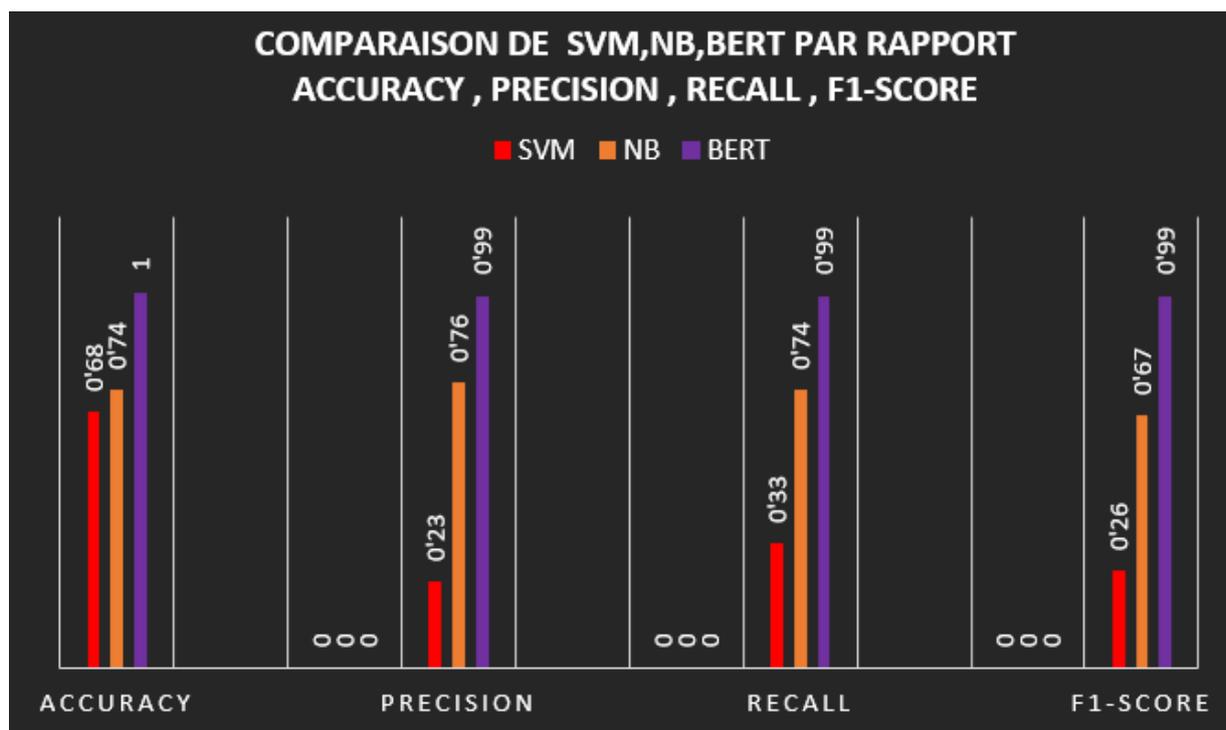


Figure III.6: Comparaison des résultats

III.8 Conclusion

Le troisième chapitre de ce mémoire a été consacré à l'implémentation de notre système. Nous avons décrit l'environnement que nous avons utilisé, ainsi que les bibliothèques employées,

les paramètres et les résultats de l'approche de gravure.

Comme nous le voyons, BERT a un excellent résultat comparé aux autres approches.

CONCLUSION GÉNÉRALE

Des millions d'étudiants s'inscrivent à des cours en ligne ouverts à la recherche d'opportunités d'apprentissage et de développement personnel. Cependant, le processus d'apprentissage implique généralement une expérience émotionnelle, qui peut affecter la participation des étudiants au cours, entraînant éventuellement un abandon en cours de route. L'engagement émotionnel décrit l'implication et l'enthousiasme des apprenants, et lorsque ces derniers sont émotionnellement engagés, ils participent plus activement aux activités d'apprentissage.

Ce travail propose une approche automatique pour détecter l'engagement émotionnel des apprenants dans les environnements d'apprentissage en ligne. Cette approche exploite l'emploi des techniques du Machine Learning dans l'analyse des avis des apprenants pour pouvoir détecter leurs émotions. Trois techniques, issues des modèles de Machine Learning, ont été testées sur une base de données qui contient les avis des apprenants. Cette base de données a été collectée à partir du site web Mark by Professor [2]. Les résultats obtenus ont montré que le modèle de BERT a prouvé sa performance dans l'analyse des émotions à travers le langage naturel.

A la fin, nous pouvons dire que les objectifs visés derrière ce travail ont été atteints. Comme perspectives de ce travail, nous proposons:

- Intégrer l'approche proposée dans une plateforme d'apprentissage en ligne, ou même l'ajouter comme Plugin dans Moodle;
- Travailler sur la construction de notre propre base de données, chose qui a été commencée en réalité au début de l'année universitaire 2022/2023;
- Explorer les forums de discussions dans la détection des émotions des apprenants

References

- [1] Yujian Cai, Xingguang Li, and Jinsong Li. Emotion recognition using different sensors, emotion models, methods and datasets: A comprehensive review. *Sensors*, 23(5):2455, 2023.
- [2] Tomáš Horváth and Imane Bouacida. Sentiment analysis and opinion mining techniques for learning analytics.
- [3] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, ukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [4] Francisca Adoma Acheampong, Henry Nunoo-Mensah, and Wenyu Chen. Transformer models for text-based emotion detection: a review of bert-based approaches. *Artificial Intelligence Review*, pages 1–41, 2021.
- [5] Le Data Scientist. À la découverte de bert, 2023.

- [6] Chang-Qin Huang, Zhong-Mei Han, Ming-Xi Li, Morris Siu-yung Jong, and Chin-Chung Tsai. Investigating students' interaction patterns and dynamic learning sentiments in online discussions. *Computers & Education*, 140:103589, 2019.
- [7] Solange Denervaud, Martina Franchini, Edouard Gentaz, and David Sander. Les émotions au cœur des processus d'apprentissage. *Revue suisse de pédagogie spécialisée*, 4:20–25, 2017.
- [8] Carla Gomes da Costa, Qin Zhou, and Aristides I Ferreira. State and trait anger predicting creative process engagement—the role of emotion regulation. *The Journal of Creative Behavior*, 54(1):5–19, 2020.
- [9] Roberta Symeonides and Carrie Childs. The personal experience of online learning: An interpretative phenomenological analysis. *Computers in Human Behavior*, 51:539–545, 2015.
- [10] Mark my professor. Mark my professor, 2010.
- [11] Linjie Zhang, Changqin Huang, Tao He, Xuemei Wu, Xizhe Wang, and Jianhui Yu. Predicting the emotional engagement in online learning: A hybrid structural equation modeling-artificial neural network approach. In *2021 International Symposium on Educational Technology (ISET)*, pages 197–201. IEEE, 2021.

- [12] Susie Lamborn, Fred Newmann, and Gary Wehlage. The significance and sources of student engagement. *Student engagement and achievement in American secondary schools*, pages 11–39, 1992.
- [13] Jennifer A Fredricks, Amy L Reschly, and Sandra L Christenson. *Handbook of student engagement interventions: Working with disengaged students*. Academic Press, 2019.
- [14] Bing Xu, Nian-Shing Chen, and Guang Chen. Effects of teacher role on student engagement in wechat-based online discussion learning. *Computers & Education*, 157:103956, 2020.
- [15] Rosalind W Picard. Toward computers that recognize and respond to user emotion. *IBM systems journal*, 39(3.4):705–719, 2000.
- [16] Satyajit Nayak, Bingi Nagesh, Aurobinda Routray, and Monalisa Sarma. A human–computer interaction framework for emotion recognition through time-series thermal video sequences. *Computers & Electrical Engineering*, 93:107280, 2021.

- [17] Xiang Feng, Yaojia Wei, Xianglin Pan, Longhui Qiu, and Yongmei Ma. Academic emotion classification and recognition method for large-scale online learning environment—based on a-cnn and lstm-att deep learning pipeline method. *International journal of environmental research and public health*, 17(6):1941, 2020.
- [18] Valentina Colonnello, Katia Mattarozzi, and Paolo M Russo. Emotion recognition in medical students: effects of facial appearance and care schema activation. *Medical Education*, 53(2):195–205, 2019.
- [19] Xiao Sun, Yezhen Song, and Meng Wang. Toward sensing emotions with deep visual analysis: a long-term psychological modeling approach. *IEEE MultiMedia*, 27(4):18–27, 2020.
- [20] Erkang Fu, Xi Li, Zhi Yao, Yuxin Ren, Yuanhao Wu, and Qiqi Fan. Personnel emotion recognition model for internet of vehicles security monitoring in community public space. *EURASIP Journal on Advances in Signal Processing*, 2021(1):1–19, 2021.
- [21] Geesung Oh, Junghwan Ryu, Euseok Jeong, Ji Hyun Yang, Sungwook Hwang, Sangho Lee, and Sejoon Lim. Drer: Deep learning–based driver’s real emotion recognizer. *Sensors*, 21(6):2166, 2021.

- [22] Regan L Mandryk, M Stella Atkins, and Kori M Inkpen. A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1027–1036, 2006.
- [23] María Malfaz and Miguel A Salichs. A new architecture for autonomous robots based on emotions. *IFAC Proceedings Volumes*, 37(8):805–809, 2004.
- [24] Michalis Feidakis, Thanasis Daradoumis, and Santi Caballé. Emotion measurement in intelligent tutoring systems: what, when and how to measure. In *2011 Third International Conference on Intelligent Networking and Collaborative Systems*, pages 807–812. IEEE, 2011.
- [25] Muhammad Anas Hasnul, Nor Azlina Ab Aziz, Salem Alelyani, Mohamed Mohana, and Azlan Abd Aziz. Electrocardiogram-based emotion recognition systems and their applications in healthcare—a review. *Sensors*, 21(15):5015, 2021.
- [26] Marlies Houben, Wim Van Den Noortgate, and Peter Kuppens. The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological bulletin*, 141(4):901, 2015.

- [27] Sonali T Saste and SM Jagdale. Emotion recognition from speech using mfcc and dwt for security system. In *2017 international conference of electronics, communication and aerospace technology (ICECA)*, volume 1, pages 701–704. IEEE, 2017.
- [28] Sebastian Zepf, Javier Hernandez, Alexander Schmitt, Wolfgang Minker, and Rosalind W Picard. Driver emotion recognition for intelligent vehicles: A survey. *ACM Computing Surveys (CSUR)*, 53(3):1–30, 2020.
- [29] Elgiz Bal, Emily Harden, Damon Lamb, Amy Vaughan Van Hecke, John W Denver, and Stephen W Porges. Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of autism and developmental disorders*, 40:358–370, 2010.
- [30] Shan Li and Weihong Deng. Deep facial expression recognition: A survey. *IEEE transactions on affective computing*, 13(3):1195–1215, 2020.
- [31] Wenjin Wang, Albertus C Den Brinker, Sander Stuijk, and Gerard De Haan. Algorithmic principles of remote ppg. *IEEE Transactions on Biomedical Engineering*, 64(7):1479–1491, 2016.

- [32] Deepak Ghimire and Joonwhoan Lee. Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines. *Sensors*, 13(6):7714–7734, 2013.
- [33] Mingmin Zhao, Fadel Adib, and Dina Katabi. Emotion recognition using wireless signals. In *Proceedings of the 22nd annual international conference on mobile computing and networking*, pages 95–108, 2016.
- [34] Jianhua Zhang, Zhong Yin, Peng Chen, and Stefano Nichele. Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion*, 59:103–126, 2020.
- [35] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine*, 29(6):82–97, 2012.
- [36] Olivier Martin, Irene Kotsia, Benoit Macq, and Ioannis Pitas. The enterface’05 audio-visual emotion database. In *22nd international conference on data engineering workshops (ICDEW’06)*, pages 8–8. IEEE, 2006.

- [37] Jinyu Li, Li Deng, Reinhold Haeb-Umbach, and Yifan Gong. Robust automatic speech recognition: a bridge to practical applications. 2015.
- [38] Björn W Schuller. Speech emotion recognition: Two decades in a nutshell, benchmarks, and ongoing trends. *Communications of the ACM*, 61(5):90–99, 2018.
- [39] Jeremy Ang, Rajdip Dhillon, Ashley Krupski, Elizabeth Shriberg, and Andreas Stolcke. Prosody-based automatic detection of annoyance and frustration in human-computer dialog. In *INTERSPEECH*, pages 2037–2040. Citeseer, 2002.
- [40] Robin Cohen. A computational theory of the function of clue words in argument understanding. In *10th International Conference on Computational Linguistics and 22nd Annual Meeting of the Association for Computational Linguistics*, pages 251–258, 1984.
- [41] Muhammad Nouman, Sui Yang Khoo, MA Parvez Mahmud, and Abbas Z Kouzani. Recent advances in contactless sensing technologies for mental health monitoring. *IEEE Internet of Things Journal*, 9(1):274–297, 2021.
- [42] Frans A Boiten. The effects of emotional behaviour on components of the respiratory cycle. *Biological psychology*, 49(1-2):29–51, 1998.

- [43] Fumihiko Yasuma and Jun-ichiro Hayano. Respiratory sinus arrhythmia: why does the heartbeat synchronize with respiratory rhythm? *Chest*, 125(2):683–690, 2004.
- [44] Changzhi Li, Julie Cummings, Jeffrey Lam, Eric Graves, and Wenhsing Wu. Radar remote monitoring of vital signs. *IEEE Microwave magazine*, 10(1):47–56, 2009.
- [45] Haobo Li, Aman Shrestha, Hadi Heidari, Julien Le Kerneec, and Francesco Fioranelli. Bi-lstm network for multimodal continuous human activity recognition and fall detection. *IEEE Sensors Journal*, 20(3):1191–1201, 2019.
- [46] William James. *The principles of psychology*, volume 1. Cosimo, Inc., 2007.
- [47] Panagiotis C Petrantonakis and Leontios J Hadjileontiadis. Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. *IEEE Transactions on affective computing*, 1(2):81–97, 2010.
- [48] Chunmei Qing, Rui Qiao, Xiangmin Xu, and Yongqiang Cheng. Interpretable emotion recognition using eeg signals. *Ieee Access*, 7:94160–94170, 2019.
- [49] Stamos Katsigiannis and Naeem Ramzan. Dreamer: A database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices. *IEEE journal of biomedical and health informatics*, 22(1):98–107, 2017.

- [50] Wanhui Wen, Guangyuan Liu, Nanpu Cheng, Jie Wei, Pengchao Shangguan, and Wenjin Huang. Emotion recognition based on multi-variant correlation of physiological signals. *IEEE Transactions on Affective Computing*, 5(2):126–140, 2014.
- [51] S Jerritta, M Murugappan, R Nagarajan, and Khairunizam Wan. Physiological signals based human emotion recognition: a review. In *2011 IEEE 7th international colloquium on signal processing and its applications*, pages 410–415. IEEE, 2011.
- [52] Azadeh Kushki, Jillian Fairley, Satyam Merja, Gillian King, and Tom Chau. Comparison of blood volume pulse and skin conductance responses to mental and affective stimuli at different anatomical sites. *Physiological measurement*, 32(10):1529, 2011.
- [53] Jia Zheng Lim, James Mountstephens, and Jason Teo. Emotion recognition using eye-tracking: taxonomy, review and current challenges. *Sensors*, 20(8):2384, 2020.
- [54] Shutao Li, Xudong Kang, Leyuan Fang, Jianwen Hu, and Haitao Yin. Pixel-level image fusion: A survey of the state of the art. *information Fusion*, 33:100–112, 2017.

- [55] Hanshu Cai, Zhidiao Qu, Zhe Li, Yi Zhang, Xiping Hu, and Bin Hu. Feature-level fusion approaches based on multimodal eeg data for depression recognition. *Information Fusion*, 59:127–138, 2020.
- [56] Zuhe Li, Yangyu Fan, Bin Jiang, Tao Lei, and Weihua Liu. A survey on sentiment analysis and opinion mining for social multimedia. *Multimedia Tools and Applications*, 78(6):6939–6967, 2019.
- [57] Binita Verma and Ramjeevan Singh Thakur. Sentiment analysis using lexicon and machine learning-based approaches: A survey. In *Proceedings of international conference on recent advancement on computer and communication*, pages 441–447. Springer, 2018.
- [58] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [59] Khin Zezawar Aung and Nyein Nyein Myo. Sentiment analysis of students’ comment using lexicon based approach. In *2017 IEEE/ACIS 16th international conference on computer and information science (ICIS)*, pages 149–154. IEEE, 2017.

- [60] Itisha Gupta and Nisheeth Joshi. Enhanced twitter sentiment analysis using hybrid approach and by accounting local contextual semantic. *Journal of intelligent systems*, 29(1):1611–1625, 2020.
- [61] Wanli Xing, Hengtao Tang, and Bo Pei. Beyond positive and negative emotions: Looking into the role of achievement emotions in discussion forums of moocs. *The Internet and Higher Education*, 43:100690, 2019.
- [62] Zachary A Pardos, Ryan SJD Baker, Maria OCZ San Pedro, Sujith M Gowda, and Supreeth M Gowda. Affective states and state tests: Investigating how affect and engagement during the school year predict end-of-year learning outcomes. *Journal of Learning Analytics*, 1(1):107–128, 2014.
- [63] Sannyuya Liu, Shiqi Liu, Zhi Liu, Xian Peng, and Zongkai Yang. Automated detection of emotional and cognitive engagement in mooc discussions to predict learning achievement. *Computers & Education*, 181:104461, 2022.
- [64] Bowen Liu, Wanli Xing, Yifang Zeng, and Yonghe Wu. Quantifying the influence of achievement emotions for student learning in moocs. *Journal of Educational Computing Research*, 59(3):429–452, 2021.

- [65] Muhammad Abbas, K Ali Memon, A Aleem Jamali, Saleemullah Memon, and Anees Ahmed. Multinomial naive bayes classification model for sentiment analysis. *IJCSNS Int. J. Comput. Sci. Netw. Secur*, 19(3):62, 2019.
- [66] J Sujanaa, S Palanivel, and M Balasubramanian. Emotion recognition using support vector machine and one-dimensional convolutional neural network. *Multimedia Tools and Applications*, 80:27171–27185, 2021.
- [67] Yukun Zhu, Ryan Kiros, Rich Zemel, Ruslan Salakhutdinov, Raquel Urtasun, Antonio Torralba, and Sanja Fidler. Aligning books and movies: Towards story-like visual explanations by watching movies and reading books. In *Proceedings of the IEEE international conference on computer vision*, pages 19–27, 2015.
- [68] Niyazi Ari and Makhamadsulton Ustazhanov. Matplotlib in python. In *2014 11th International Conference on Electronics, Computer and Computation (ICECCO)*, pages 1–6. IEEE, 2014.
- [69] Meng Wang and Fanghui Hu. The application of nltk library for python natural language processing in corpus research. *Theory and Practice in Language Studies*, 11(9):1041–1049, 2021.

- [70] Bo Pang, Erik Nijkamp, and Ying Nian Wu. Deep learning with tensorflow: A review. *Journal of Educational and Behavioral Statistics*, 45(2):227–248, 2020.
- [71] MyGreatLearning. Python numpy tutorial: Learn numpy arrays with examples, 2023.