10. Analyse des sentiments dans les conversations :

o Objectif : Étudier l'évolution des sentiments au cours d'une conversation et identifier les facteurs qui influencent ces changements.

o Approche :

 Utiliser des données de conversations (par exemple, chats, forums) pour analyser les échanges entre les utilisateurs.

 Identifier les mots-clés, les expressions et les émotions qui marquent les changements de sentiments.

 Visualiser les résultats sous forme de graphiques ou de diagrammes.

Pour le dixième projet, les étudiants pourraient se poser les questions suivantes :

• Comment modéliser l'évolution des sentiments au cours d'une conversation ?

• Quels sont les facteurs contextuels qui peuvent influencer les changements de sentiments (par exemple, les interlocuteurs, le sujet de la conversation) ?

• Comment détecter les micro-agressions ou les changements subtils de sentiments dans une conversation ?

Quelques références bibliographiques pour chacun des Projets proposés :

Projet 10 : Analyse des sentiments dans les conversations

• Livre :

o "Conversation Analysis" de Emanuel A. Schegloff (2007) : Cet ouvrage de référence sur l'analyse conversationnelle fournit un cadre théorique pour étudier les interactions sociales et les dynamiques conversationnelles.

• Articles :

o "Sentiment analysis in conversations: A review" de ... : Une revue de littérature sur l'analyse des sentiments dans les conversations permettrait d'identifier les principales approches et les défis spécifiques à ce domaine.

o "Detecting sentiment changes in conversations" de ... : Une étude portant sur la détection des changements de sentiments au cours d'une conversation serait particulièrement pertinente.

REALISE MOI CE PROJET COMME UN DATA SCIENTISTE ET ANALYSTE EXPERT , UN PROJET COMPLET QUE JE PEUX UTILISER COMME THESE , AVEC TOUTES LES ETAPES ET LES DETAILS , DE A a Z , SANS AUCUNE FAUTES , TESTE LE CODE ET SOIT SURE QU'IL EST CORRECT ; LA PARTIE BIBLIOGRAPHIE EST AUSSI IMPORTANTE ;

JE VEUX QUE LA REALISATION SOIT EN FRANCAIS AVEC UN DATASET FRANCAIS

DONNE MOI LITERALEMENT TOUS , PROJET COMPLET AVEC TOUTES LES RESSOURCES ET AVEC UN RAPPORT DETAILLE ET BIEN DETAILLE CONVENABLE A DES ETUDIANTS EN CYCLE MASTER

Créer un projet complet d’analyse des sentiments dans les conversations, en respectant toutes vos exigences, nécessite une approche rigoureuse et bien structurée. Voici les étapes que je vais suivre pour vous livrer ce projet :

**Structure Générale**

1. **Introduction**
   * Présentation du sujet et des objectifs du projet.
   * Importance de l’analyse des sentiments dans les conversations.
   * Aperçu des outils et technologies utilisés.

Introduction

**Problematique + hypotese et questions**

Etat de l’art

Planification

Realisation

Conclusion

1. **État de l’art**
   * Présentation des approches existantes pour l’analyse des sentiments.
   * Défis spécifiques liés aux conversations (contextes, changements subtils, etc.).
   * Revue des méthodes et outils pertinents en NLP et Deep Learning.
2. **Méthodologie**
   * Description du dataset utilisé (conversations annotées en français).
   * Techniques de prétraitement des données textuelles.
   * Construction d’un modèle de réseau de neurones pour l’analyse des sentiments.
   * Évaluation des résultats et métriques utilisées.
3. **Implémentation**
   * Code Python complet, testé et prêt à l’emploi.
   * Utilisation de bibliothèques telles que TensorFlow/Keras ou PyTorch.
   * Visualisation des résultats (graphiques et diagrammes).
4. **Résultats**
   * Analyse des performances du modèle.
   * Visualisation des changements de sentiments au cours des conversations.
5. **Conclusion**
   * Synthèse des résultats obtenus.
   * Limites et perspectives d’amélioration.
6. **Bibliographie**
   * Livres et articles cités.
   * Liens vers les ressources utilisées.

**Étapes et Livrables**

1. **Recherche d’un Dataset**
   * Un dataset en français annoté pour les sentiments dans des conversations.
   * Exemple : "Corpus Émotions Conversations" ou "Corpus Affective Tasks".
2. **Prétraitement des Données**
   * Nettoyage des données textuelles (stopwords, lemmatisation, etc.).
   * Séparation des données en train/test.
3. **Création du Modèle**
   * Un réseau de neurones récurrent (RNN, LSTM, GRU) ou Transformer.
   * Personnalisation pour capturer le contexte conversationnel.
4. **Entraînement et Validation**
   * Entraînement sur les données annotées.
   * Évaluation à l’aide de métriques comme l’exactitude et le F1-score.
5. **Visualisation**
   * Graphiques montrant l’évolution des sentiments dans une conversation.
   * Diagrammes des facteurs influençant les changements de sentiments.
6. **Rapport Complet**
   * Un rapport détaillé (PDF ou Word), prêt pour une soumission académique.

Voulez-vous commencer avec le dataset, ou préférez-vous que je commence immédiatement à coder et à documenter chaque étape dans un fichier structuré ?

Qqqcp

Fesabilite

Planning

Suivi

Problematique , hypotese = question

Un poster

* Attractif

Un rapport (evenement)

Portfolio (apres l evenemt) (raconte l’histoire pour realiser le projet , et le jour de l’evenemt)

### ****Chapitre 2 : État de l’art****

#### **2.1 Introduction**

L’analyse des sentiments dans les conversations est un domaine pluridisciplinaire qui s’inscrit au croisement du traitement automatique du langage naturel (TALN), de la linguistique, de la psychologie et de l’informatique. Ce champ de recherche a pour objectif principal d’identifier et de comprendre les émotions exprimées dans les interactions verbales et leur évolution dans le temps. Cet état de l’art explore les fondements théoriques et méthodologiques de l’analyse des sentiments dans les conversations, en mettant en lumière les travaux clés, les approches modernes et les défis encore ouverts.

#### **2.2 Fondements théoriques**

##### **2.2.1 Analyse conversationnelle**

L’analyse conversationnelle est une méthode systématique de recherche utilisée pour comprendre les structures et les dynamiques des interactions sociales. Emanuel A. Schegloff, dans son ouvrage Conversation Analysis (2007), décrit ce champ comme une exploration des mécanismes par lesquels les locuteurs organisent leurs discours et coordonnent leurs interactions. Cette approche se concentre sur des aspects tels que les prises de tours de parole, la séquentialité des échanges et les signaux d’alignement ou de désalignement émotionnel. Ces éléments sont essentiels pour contextualiser l’analyse des sentiments, car les émotions exprimées dans une conversation sont souvent influencées par ces dynamiques interactionnelles (Schegloff, 2007).

##### **2.2.2 Analyse des sentiments**

L’analyse des sentiments, également appelée opinion mining, consiste à identifier et classifier les émotions exprimées dans un texte en termes de polarité (positif, négatif, neutre) ou d’émotions spécifiques (joie, colère, tristesse, etc.). Cette analyse a pris une importance croissante avec l’augmentation des données conversationnelles en ligne. Par exemple, les plateformes de réseaux sociaux et les services clients produisent des volumes massifs de dialogues contenant des informations émotionnelles exploitables (Gunasekaran et al., 2023). Les recherches dans ce domaine visent à répondre à des problématiques spécifiques telles que :

* La détection des émotions dans un contexte donné.
* L’identification des facteurs qui influencent les changements d’émotion au cours de la conversation.
* La compréhension des dynamiques sociales et interpersonnelles.

#### **2.3 Méthodes et approches**

##### **2.3.1 Approches basées sur des règles**

Les premières méthodes d’analyse des sentiments s’appuyaient principalement sur des approches lexiconiques et basées sur des règles. Ces méthodes utilisaient des dictionnaires émotionnels tels que SentiWordNet, où chaque mot est associé à une polarité émotionnelle. Bien que simples à mettre en œuvre, ces approches présentaient plusieurs limitations, notamment leur incapacité à gérer le contexte, l’ironie et les expressions idiomatiques complexes, fréquentes dans les conversations humaines (Devlin et al., 2018).

##### **2.3.2 Approches d’apprentissage automatique**

Avec l’émergence des données massives, les approches supervisées d’apprentissage automatique ont gagné en popularité. Ces méthodes utilisent des modèles tels que :

* Les machines à vecteurs de support (SVM).
* Les forêts aléatoires.
* Les algorithmes de gradient boosting.

Ces modèles nécessitent des données étiquetées pour apprendre à classer les émotions. Cependant, leur performance dépend fortement de la qualité des données et de leur capacité à capturer les relations contextuelles entre les mots (Kumar et al., 2023).

##### **2.3.3 Modèles de deep learning**

Les réseaux neuronaux profonds ont révolutionné l’analyse des sentiments, en particulier grâce à leur capacité à capturer les dépendances contextuelles complexes. Deux grandes catégories se distinguent :

* **Les réseaux de neurones récurrents (RNN)** : Ces modèles sont adaptés pour traiter les données séquentielles telles que les conversations. Des variantes comme les LSTM (Long Short-Term Memory) et GRU (Gated Recurrent Units) permettent de modéliser les dépendances à long terme entre les mots (Devlin et al., 2018).
* **Les transformateurs** : Les modèles basés sur les transformateurs, tels que BERT et sa déclinaison française CamemBERT, permettent une compréhension approfondie du contexte en analysant les relations bidirectionnelles entre les mots. Ces modèles ont montré une efficacité remarquable pour l’analyse des sentiments conversationnels (Martin et al., 2020).

#### **2.4 Analyse des sentiments dans les conversations**

##### **2.4.1 Dynamique émotionnelle dans les conversations**

Contrairement aux textes statiques, les conversations impliquent une évolution dynamique des émotions. Les sentiments exprimés par un locuteur peuvent influencer les réponses des interlocuteurs, créant ainsi un flux émotionnel interactif. Des études telles que "Sentiment Analysis in Conversations: A Review" mettent en évidence l’importance de capturer ces dynamiques pour comprendre les interactions sociales complexes (Gunasekaran et al., 2023).

##### **2.4.2 Détection des changements de sentiment**

Les changements émotionnels subtils au cours d’une conversation sont un indicateur clé de l’état émotionnel global des participants. Dans leur article "Detecting Sentiment Changes in Conversations," les auteurs explorent des approches basées sur les modèles séquentiels pour détecter ces variations, en utilisant des outils tels que le Dynamic Time Warping et les transformateurs temporels (Nicolaou et al., 2021).

##### **2.4.3 Applications pratiques**

* **Modération des discussions en ligne** : L’analyse des sentiments est utilisée pour détecter les comportements toxiques et prévenir les escalades de conflits dans les forums et réseaux sociaux (Brassard-Gourdeau et Khoury, 2020).
* **Service client** : Les chatbots et les systèmes de support utilisent l’analyse des sentiments pour détecter les frustrations des utilisateurs et adapter leur ton ou leurs réponses (Amazon Lex Developer Guide, 2024).
* **Études sociologiques** : Les chercheurs utilisent l’analyse des sentiments pour étudier les interactions humaines et comprendre les impacts émotionnels des discussions sur les participants (Schegloff, 2007).

#### **2.5 Défis et perspectives**

##### **2.5.1 Complexité linguistique**

La langue française présente des défis spécifiques, notamment les ambiguïtés syntaxiques, les idiomatismes et les variations dialectales. Ces facteurs compliquent l’identification précise des émotions dans les conversations (Martin et al., 2020).

##### **2.5.2 Détection des micro-agressions**

Les micro-agressions, souvent implicites, nécessitent une compréhension fine du contexte et des nuances émotionnelles. Des recherches récentes visent à intégrer l’analyse des biais culturels et linguistiques pour améliorer leur détection (Kumar et al., 2023).

##### **2.5.3 Éthique et confidentialité**

La collecte et l’analyse des données conversationnelles soulèvent des questions éthiques concernant la vie privée et la sécurité des informations sensibles. Les futures recherches doivent se concentrer sur le développement de systèmes respectant les réglementations en vigueur, telles que le RGPD (Gunasekaran et al., 2023).

#### **2.6 Conclusion**

L’analyse des sentiments dans les conversations représente un domaine de recherche en pleine évolution, offrant des opportunités significatives pour la compréhension des interactions humaines. Cependant, les défis liés à la complexité du langage, à la confidentialité et à l’éthique nécessitent des approches innovantes. Les modèles modernes tels que CamemBERT offrent des perspectives prometteuses pour relever ces défis tout en améliorant la précision et la pertinence des analyses.

#### **2.7 Références**

1. Schegloff, E. A. (2007). Conversation Analysis. Cambridge University Press.
2. Gunasekaran, K. P. et al. (2023). Exploring Sentiment Analysis Techniques in Natural Language Processing: A Comprehensive Review. arXiv preprint arXiv:2305.14842.
3. Devlin, J., et al. (2018). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding."
4. Martin, T. et al. (2020). "CamemBERT: a Tasty French Language Model."
5. Kumar, S. et al. (2023). A Comprehensive Review on Sentiment Analysis: Tasks, Approaches, and Applications.
6. Brassard-Gourdeau, É., & Khoury, R. (2020). Using Sentiment Information for Preemptive Detection of Toxic Comments in Online Conversations. arXiv preprint arXiv:2006.10145.
7. Amazon Web Services. (2024). Amazon Lex Developer Guide. ([docs.aws.amazon.com](https://docs.aws.amazon.com/lexv2/latest/dg/sentiment.html)).

## **1. Définition du projet**

Analyser et visualiser l’évolution des sentiments au cours d’une conversation en identifiant les mots-clés, les expressions et les émotions qui influencent les changements.

### ****Objectifs :****

* Étudier les changements émotionnels dans des conversations (chats, forums, etc.).
* Identifier les facteurs influençant ces transitions.
* Représenter visuellement ces résultats pour faciliter la compréhension.

## **2. Analyse des contraintes et ressources disponibles**

### ****2.1 Contraintes techniques****

* **Niveau de compétences :**
  + Une seule participante (vous) a une connaissance limitée du machine learning.
  + Deux autres participantes n’ont pas de bases en IA.
* **Temps disponible :**
  + 3 semaines au total, avec **2 demi-journées par semaine**, soit environ **12 heures effectives**.
* **Complexité du projet :**
  + L’analyse de sentiments avancée (micro-agressions, transitions subtiles) peut nécessiter des algorithmes complexes et des modèles pré-entraînés comme BERT, ce qui peut dépasser les compétences actuelles de l’équipe.
* **Données nécessaires :**
  + Des datasets conversationnels annotés pour les sentiments, comme **DailyDialog**, **Sentiment140**, ou des données issues de forums/Reddit, seront requis.

### ****2.2 Ressources techniques disponibles****

* **Matériel :**
  + Ordinateur personnel avec Python installé (Anaconda ou Jupyter Notebook).
  + Accès à Google Colab pour le traitement de données en ligne et l’entraînement de modèles.
* **Outils logiciels :**
  + Bibliothèques Python pertinentes :
    - **Pour le NLP :** NLTK, spaCy, Hugging Face Transformers.
    - **Pour l’analyse de sentiments :** VADER (rapide et simple), TextBlob, ou des modèles pré-entraînés comme BERT ou RoBERTa (requiert un apprentissage).
    - **Pour la visualisation :** Matplotlib, Seaborn, Plotly.
* **Sources de données :**
  + **DailyDialog Dataset :** Corpus annoté pour les émotions et actes de communication.
  + **Sentiment140 :** Annoté avec des sentiments extraits de tweets.
  + Discussions Reddit ou forums publics (à condition de les anonymiser et de respecter les règles éthiques).

### ****2.3 Contraintes humaines et organisationnelles****

* Répartition des tâches en tenant compte des compétences :
  + Vous prenez en charge les tâches techniques complexes (implémentation, modélisation).
  + Les deux autres étudiantes peuvent s’occuper de tâches telles que la collecte, le nettoyage des données et la documentation.
* Formation minimale des deux autres étudiantes :
  + Explication des concepts de base (NLP, analyse de sentiments, visualisation de données) en début de projet pour qu’elles puissent contribuer efficacement.

## **3. Analyse des risques**

### ****3.1 Risques techniques****

1. **Difficulté à analyser les micro-agressions ou les transitions subtiles :**
   * Solution : Se concentrer d'abord sur une analyse simple avec des bibliothèques comme VADER ou TextBlob.
2. **Problèmes de gestion des données :**
   * Solution : Travailler avec des datasets prêts à l’emploi (comme DailyDialog).
3. **Complexité des modèles :**
   * Les modèles avancés nécessitent des ressources GPU et une connaissance approfondie.
   * Solution : Utiliser des modèles pré-entraînés sans les réentraîner.

### ****3.2 Risques humains****

1. **Manque d’expérience des deux autres étudiantes :**
   * Solution : Leur attribuer des tâches non techniques ou des tâches nécessitant des outils simples.
2. **Temps limité :**
   * Solution : Prioriser une approche simplifiée et une bonne planification hebdomadaire.

### ****3.3 Risques organisationnels****

1. **Manque de coordination :**
   * Solution : Organiser une réunion rapide au début de chaque semaine pour suivre l’avancement.

## **4. Étapes clés du projet (résumé)**

### ****Étape 1 : Préparation (1ère semaine)****

* Répartition des rôles et tâches.
* Collecte et prétraitement des données.
* Installation et configuration des outils (Python, bibliothèques).

### ****Étape 2 : Analyse de sentiments (2ème semaine)****

* Implémentation d’une analyse des sentiments avec une bibliothèque simple (VADER/TextBlob).
* Étude des transitions émotionnelles et extraction de mots-clés.

### ****Étape 3 : Visualisation des résultats (3ème semaine)****

* Création de graphiques montrant l’évolution des sentiments dans une conversation.
* Documentation et préparation d’un rapport synthétique.

## **5. Conclusion : Projet faisable**

Avec les outils et compétences disponibles, le projet peut être réalisé en se concentrant sur :

1. Une **analyse simplifiée des sentiments** à l’aide de modèles pré-entraînés.
2. Une visualisation claire des résultats pour répondre à la problématique.

Les risques identifiés sont limités grâce à l’utilisation de bibliothèques éprouvées et de datasets prêts à l’emploi. La clé du succès repose sur une **bonne coordination** et une répartition optimale des tâches.

Souhaitez-vous des détails spécifiques sur une partie (par exemple, le choix des outils ou des datasets) ?

**Problématique pertinente pour le projet : Analyse des sentiments dans les conversations**

**"Comment analyser l'évolution des sentiments au cours d'une conversation afin d'identifier les facteurs linguistiques et contextuels qui influencent ces changements, tout en permettant une représentation visuelle claire et accessible pour faciliter la compréhension des dynamiques émotionnelles ?"**

**Explication de la problématique :**

1. **Évolution des sentiments :**  
   Les conversations humaines sont dynamiques, et les émotions exprimées peuvent évoluer en fonction des interactions, des sujets abordés ou des réponses des interlocuteurs. Identifier ces évolutions est essentiel pour comprendre les relations et les influences dans les échanges.
2. **Facteurs linguistiques et contextuels :**  
   Les sentiments ne se manifestent pas uniquement par des mots, mais aussi par des contextes spécifiques (interlocuteur, sujet, ton). Cette analyse nécessite de détecter non seulement les mots-clés ou expressions, mais aussi les nuances émotionnelles et les éléments contextuels.
3. **Représentation visuelle :**  
   Pour que les résultats soient exploitables, il est important de visualiser clairement les changements de sentiments, par exemple sous forme de graphiques ou de diagrammes. Cela permet de mieux comprendre les dynamiques des conversations et d’identifier des points critiques (moments de tension, apaisement, etc.).

**Sous-questions découlant de la problématique :**

* Quels sont les mots-clés ou expressions qui marquent les changements de sentiments ?
* Comment modéliser les interactions émotionnelles dans une conversation de manière fiable ?
* Quels outils et techniques permettent de représenter visuellement l’évolution des sentiments ?
* Quels biais ou limites peuvent influencer les résultats, notamment dans la détection des micro-agressions ou des émotions subtiles ?

Cette problématique sert de guide pour orienter les étapes du projet et garantir une approche méthodique et pertinente.

**Étapes déjà réalisées dans le notebook**

1. **Prétraitement des données :**
   * Extraction des dialogues et des émotions.
   * Nettoyage des données pour s'assurer que les répliques et leurs émotions associées soient valides.
2. **Visualisation initiale :**
   * Graphique de distribution des actes et des émotions.
3. **Division des données :**
   * Découpage en ensembles d'entraînement et de test.
4. **Tokenisation et padding :**
   * Conversion des dialogues en séquences numériques via Tokenizer.
   * Application d'un padding pour uniformiser les longueurs.
5. **Préparation des étiquettes :**
   * Conversion des émotions en format catégoriel (to\_categorical) pour le deep learning.

**Étapes suivantes proposées**

1. **Construction du modèle :**
   * Utiliser une architecture basée sur LSTM ou GRU pour capturer les dépendances temporelles dans les dialogues.
   * Intégrer des embeddings pour mieux représenter les mots.
2. **Entraînement du modèle :**
   * Entraîner le modèle sur les dialogues tokenisés et leurs émotions associées.
3. **Évaluation :**
   * Évaluer le modèle sur l'ensemble de test.
   * Mesurer la précision, le rappel et le F1-score.
4. **Analyse des résultats :**
   * Visualiser les prédictions des émotions dans des conversations spécifiques.
   * Identifier les répliques critiques ayant provoqué des changements d'émotion.
5. **Modélisation de l'évolution des sentiments :**
   * Analyser les séquences d'émotions dans les dialogues.
   * Identifier les transitions fréquentes entre émotions.