**Rapport d'Analyse des Sentiments sur les Avis de Travel Insured International**

## **1. Introduction**

### **1.1 Contexte**

L'analyse des avis clients est cruciale pour les entreprises cherchant à améliorer leurs services. Ce projet vise à analyser les sentiments des utilisateurs à partir des avis Trustpilot concernant Travel Insured International.

### **1.2 Objectifs**

* Scraper les avis clients depuis Trustpilot
* Analyser les sentiments avec un modèle NLP (BERT/DistilBERT)
* Développer une API pour le traitement des avis
* Créer une interface interactive pour la visualisation des tendances

## **2. Architecture et Modèle NLP Utilisé**

### **2.1 Collecte des Données**

* Scraping effectué avec **BeautifulSoup**
* Extraction de 900 avis
* Colonnes principales : **rate(1-5), title, content, review\_date**

### **2.2 Prétraitement des Données**

* Création de la colonne **Reviews** représentant les avis à partir de **title** et **content**
* Nettoyage des avis (transformation du texte, stopwords, lemmatisation)
* Tokenisation et création de la colonne **tokenized\_reviews**

### **2.3 Modèle NLP Utilisé**

* **DistilBERT** pré-entraîné sur des corpus de sentiment analysis
* Classification des avis en **positif et négatif**
* API d'inférence basée sur **Hugging Face**
* Mapping des avis analysés avec leur sentiment (colonne **Sentiment**)

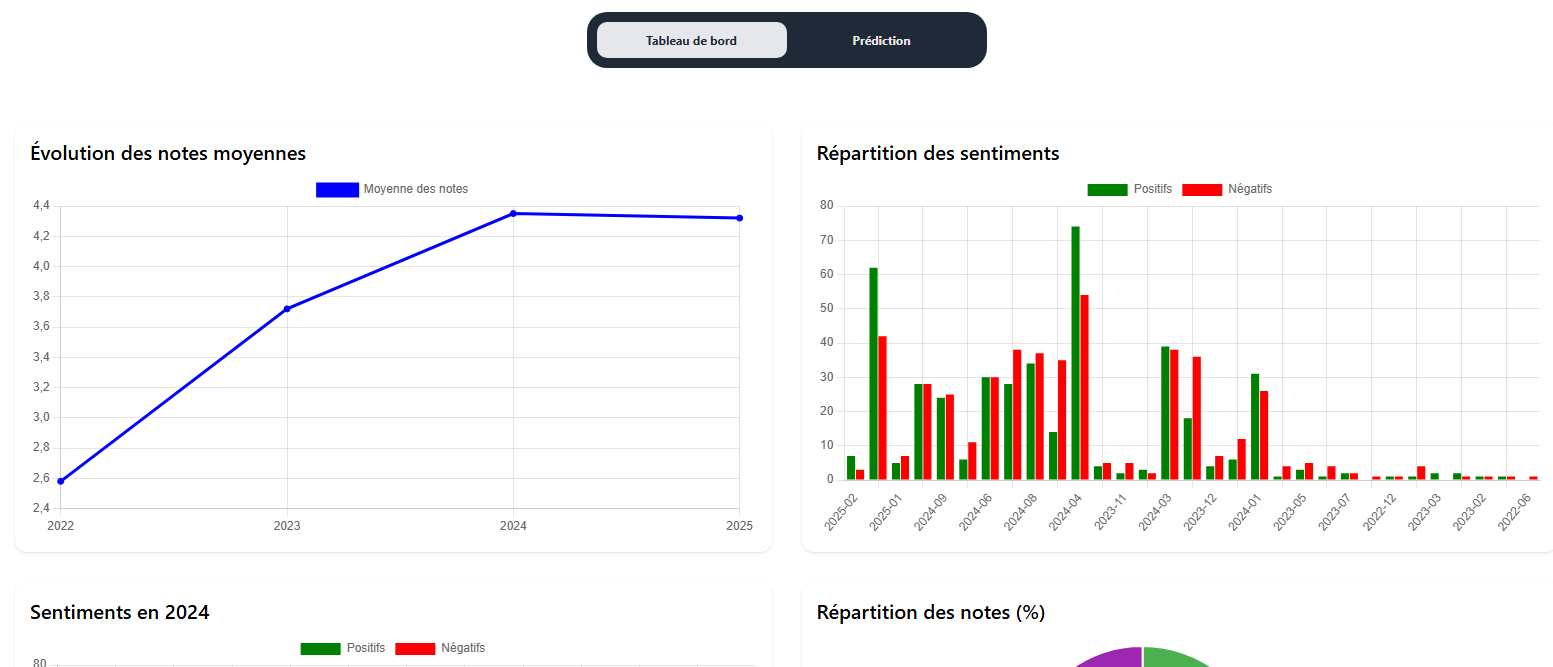
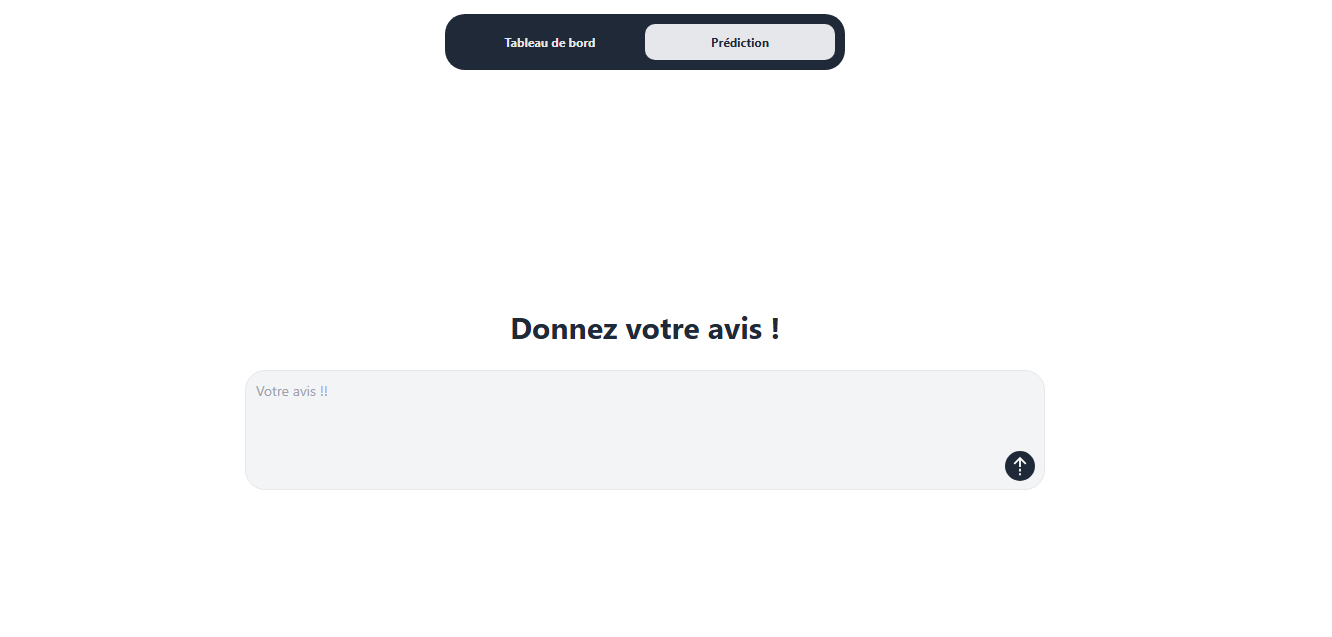
## **3. Développement de l'API et Interface**

### **3.1 API d'Inférence**

* Développée avec **Flask**
* Exposition d'un endpoint permettant de classifier un avis en temps réel

### 

### **3.2 Interface Utilisateur**

* Interface web en **JavaScript** pour l'affichage des résultats et tendances
* Interface de prédiction de sentiment d’avis 

## **4. Conclusion et Perspectives**

### **4.1 Conclusion**

L'analyse des sentiments appliquée aux avis de Travel Insured International a permis d'identifier les tendances générales des clients. L'utilisation de DistilBERT a offert des résultats précis et pertinents.

### **4.2 Perspectives**

* Intégration d'autres facteurs influents (météo, tarifs, événements économiques, situations géographique)
* Optimisation du modèle par fine-tuning sur l’ensemble des avis collectés
* Une feature pipeline pour collecter les avis en temps réel