Leveraging NLP for Market Risk Assessment and Decision Support

Theresa Sumarta
UC Berkeley
04/16/2025

Introduction

Problem

Market-entry decisions carry risks: political, economic, technological, sector-specific.

Traditional tools = static reports & expert opinions → lack depth & real-world context.

Current models underuse open-source data like news articles.

Solution

A **proof-of-concept framework** using *The New York Times* (2010–2020) articles.

Employs **transformer-based models** for:

- Risk classification (Economic, Technological, Business, etc.
- Risk-focused question answering (QA) via **RAG**.

Dataset

Data Source

Based on **N24News Corpus** (Zhu et al., 2021)

Derived from *The New York Times (2010-2020)*

60,000 articles → **Filtered to 12,031** relevant to market-entry risk

Category Focus

Selected **6** risk-relevant categories:

- Economy
- Technology
- Global Business
- Real Estate
- Your Money
- Automobiles

Processing Pipeline

Original dataset provided headlines, abstracts, and article URLs.

Used **newspaper3k** to extract **full article text**

Preprocessing: tokenization, sentence segmentation, HTML cleaning

Methodology: Risk Level Classification

Manual Annotation

690 articles manually annotated with risk level (**Low, Medium or High**)

Fine-tuned 6 models

BERT, DeBERTa-v3, T5, RoBERTa, DistilBERT, FinBERT

Stratified **70/15/15** train-validation-test split

Trained for **5 epochs** with optimized hyperparameters

Model Evaluation

Metrics: precision, recall, F1, and accuracy

Label rest of dataset

Best model used to label remaining **11,341 articles**

Random samples manually reviewed for **quality assurance**

Methodology: RAG-based QA Framework

01 **FAISS vector store**

With multi-qa-mpnet-base-dot-v1 embeddings
Retrieved **top 5** relevant articles per query through similarity search

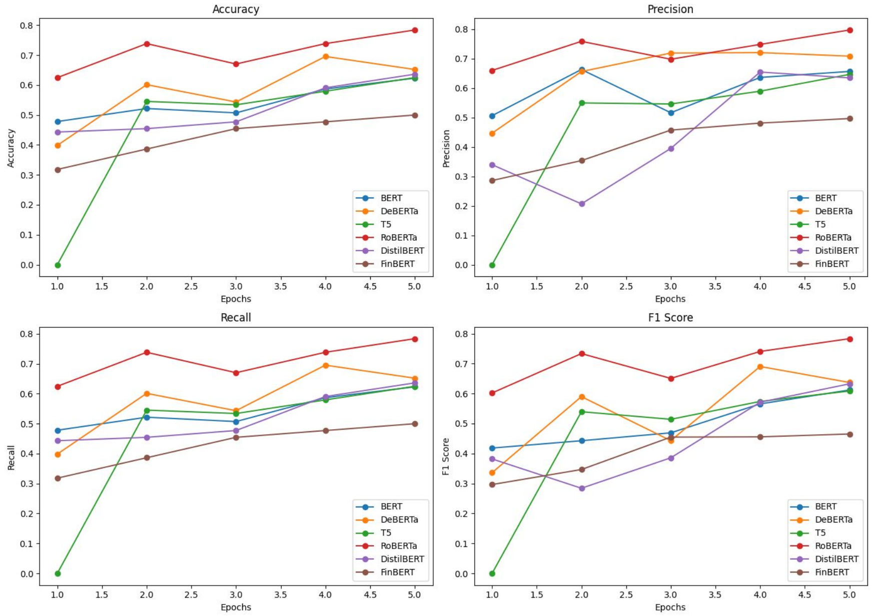
02 **Fine-tuned Models**

GPT-2, Mistral, and LLaMA 3.1 were fine-tuned for response generation
Applied **prompt engineering**

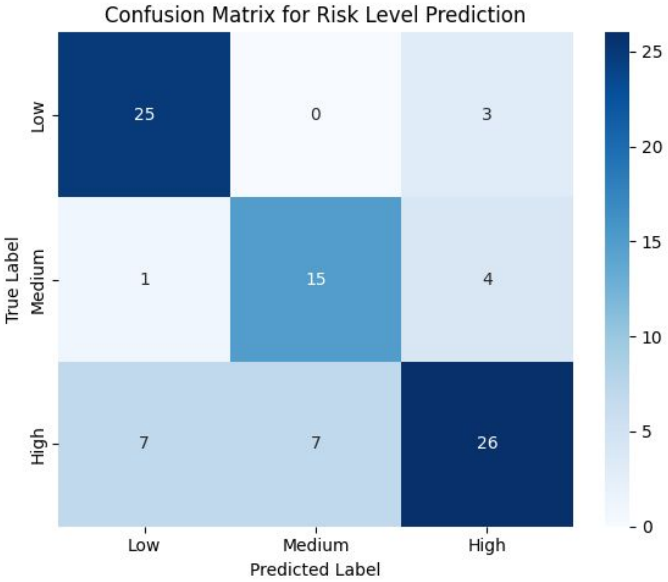
03 **Model Evaluation**

Evaluated with metrics such as **F1, ROUGE-L, METEOR, perplexity**, and **faithfulness**

Results: Risk Level Classification



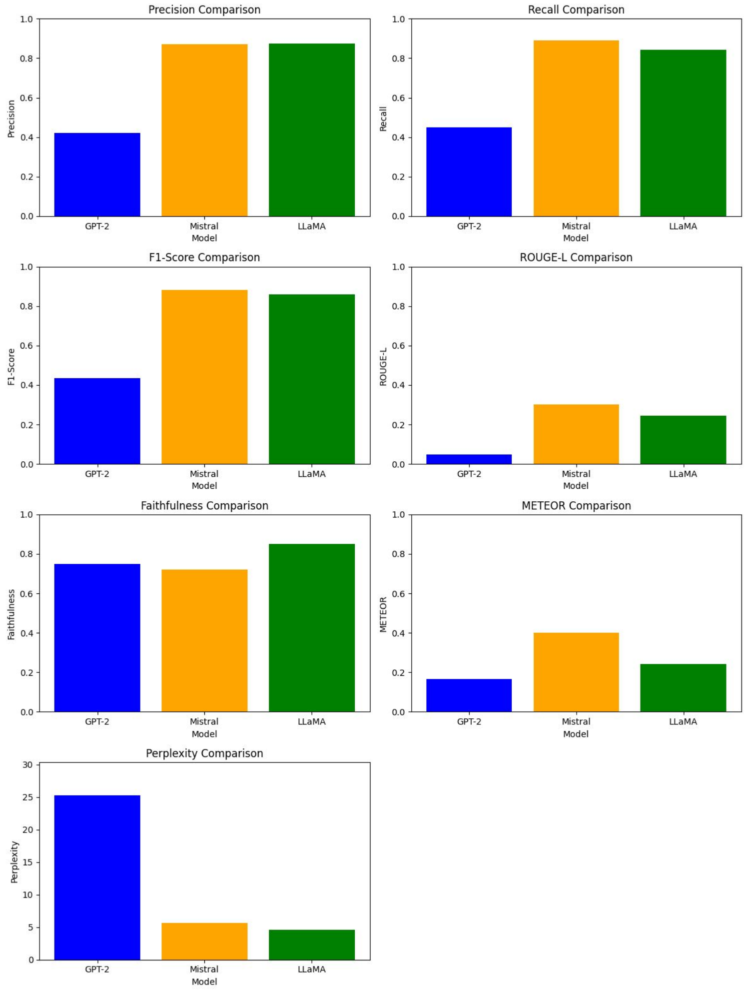
The optimized **RoBERTa model** outperformed all other models, achieving an accuracy of 0.75, precision of 0.754, recall of 0.75, and F1-score of 0.747.



High degree of class separability

Results:
Question Answering

- **LLaMA 3.1** achieved highest **precision (0.875)** and **faithfulness (0.851)**
- Best configuration: temp = 0.7, top-p = 0.95, repetition penalty = 1.0
- **Mistral** slightly better in fluency and semantic relevance
- **GPT-2** scored lowest across all evaluation metrics



Qualitative Comparison

"Which economic risks have been reported for electric vehicle companies entering the US market?'

Mistral

Generic risks

Output: Risk Type: Economic

Risk Level: High/Medium **Inconsistent labeling**

1) Dependence on Electricity Sources - The impact of electric vehicles (EVs) on reducing greenhouse gas emissions largely depends upon the region's electrical grid composition. In regions dominated by fossil fuels such as coal, EV charging might result in similar levels of CO_2 emission compared to conventional internal combustion engine vehicles. This uncertainty could negatively affect consumer perception towards EV brands and potentially limit growth opportunities.

(Referenced Risks: Environmental Regulations & Compliance; Energy Prices)

LLaMA 3.1

1. **Competition from established hybrid models**: The report highlights the growing popularity of hybrid models, which could lead to increased competition for electric vehicle companies.

Risk Type: Industry

Risk Level: Medium

Distinct economic risk

2. **Volatility in energy costs**: The report notes that the environmental impact of electric vehicles depends on the source of the electricity used to charge them, which can vary significantly depending on the region.

Risk Type: Industry

Risk Level: Medium **Clear risk level**

Conclusion

- **RoBERTa** excelled in risk classification
- **LLaMA 3.1** performed best in QA generation
- Demonstrated feasibility of **automated risk assessment**
- Supports **informed decision-making** for market entry
- Foundation for future **scalable, real-time tools**

Thank You