The NLP Landscape in Finance

Opportunities and Challenges

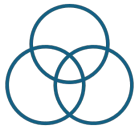
Suzet Nkwaya
Data Scientist, MFS

November 13th, 2024

Agenda

Overview of NLP in Finance

Key NLP Tasks



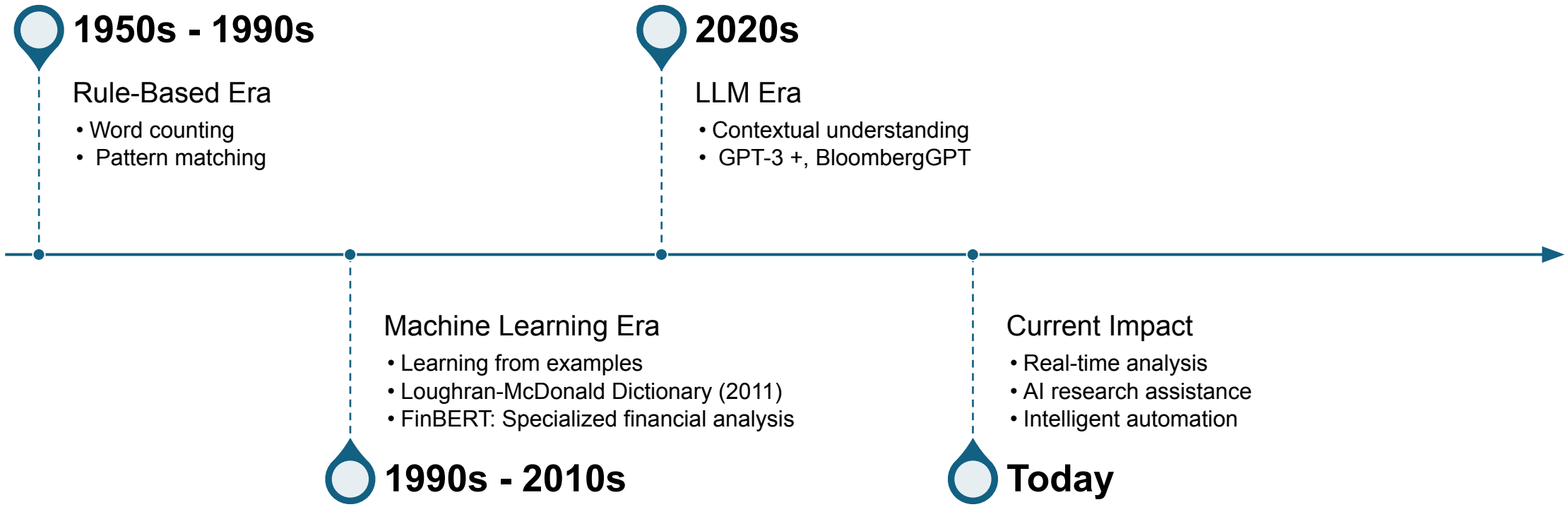
Traditional NLP and GenAI

Use Cases and Practical Examples

Conclusion

The Evolution of NLP

Rule-Based Systems to LLMs



Core NLP Tasks in Finance

Classification & NER

Identify and classify entities with specific meanings from text, such as names, places, organizations,
SEC filings, news categorization

Sentiment Analysis

Analyzing market sentiment from text sources

Earnings calls, social media, news

Text Summarization

Creating concise summaries of lengthy documents

Research reports, regulatory filings

Text Generation

Creating financial content and reports

Automated reporting, research summaries

Textual Classification

Industry

Groups companies based on activities and performance

Topic

Sorts financial documents into themes

ESG

Identifies ESG-related data

Performance metrics for similarity and return attribution

Method	Avg Pairwise Correlation				R^2
	$k = dynamic$	$k = 1$	$k = 5$	$k = 10$	
GICS Sector	0.362	-	-	-	0.052
GICS Industry	0.409	-	-	-	0.106
BERT-FT	-	0.450	0.430	0.421	0.100
LF-FT-512	-	0.449	0.425	0.415	0.110
LF-FT-1024	-	0.319	0.312	0.309	0.096
SBERT-FT	-	0.458	0.438	0.428	0.100
SBERT-PT	-	**0.471**	**0.443**	**0.432**	0.114
GPT-ada	-	0.462	0.438	0.425	0.117
PaLM-gecko	-	**0.471**	0.442	0.431	**0.119**

Source: Vamvourellis, et.al (2023) arXiv:2308.08031

Name Entity Recognition (NER)

NER in Finance

Identifies critical elements:

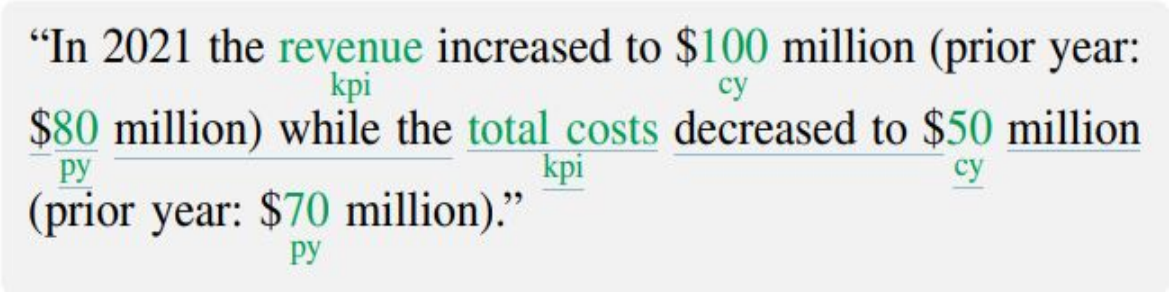
- company names
- stock ticker
- financial indicators

foundation for many applications

- Fraud Detection
- Regulatory Compliance
- Sentiment Analysis

The Evolution of NER

- Rule-based methods (first generation)
- Machine learning (supervised and unsupervised)
- LLMs



Source: Hillebrand, et.al. (2022)
arXiv:2208.02140

Sentiment Analysis

From Words to Context

Traditional Methods (Pre-2010)

- **Loughran-McDonald Dictionary: Finance-specific lexicon**
  - Recognized financial context (e.g., "liability," "expense")
  - Limited by rigid word categorization

```
if "strong growth" in earnings_call:
sentiment = "positive"
# If only it were this simple!
```

Example : *"Our aggressive debt restructuring led to outstanding growth"*

Dictionary Analysis:

- 🔴 "aggressive" - flagged as negative (risk indicator)
- 🔴 "debt" - flagged as negative (liability)
- 🟢 "outstanding" - flagged as positive
- 🟢 "growth" - flagged as positive

- **Machine Learning Breakthrough**
  - SVM, RF, XGBoost, etc
  - Pattern recognition in large datasets
  - Challenge: Domain adaptation
- **Word Embeddings Evolution**
  - Word2Vec, GloVe, FastText
  - Enhanced semantic understanding

The LLM Revolution

- **Comprehensive Analysis**
  - Decodes subtle indicators (e.g., earnings calls)
  - Understands specialized terminology
- **Multimodal Processing**
  - Text, audio, video integration
  - Resilience against misinformation

GPT vs FinBERT

Methodology

- **Zero-shot**: Models tested without specific examples
- **Few-shot**: Models provided with similar example cases
- **Fine-tuning**: Domain-specific training on financial data
- **Dataset**: Financial PhraseBank
  - 4,845 financial news sentences from LexisNexis
  - Annotated by 16 finance experts for stock price impact
  - Split: 60% training
  - 20% validation, 20% test
  - 10-fold cross-validation for result verification

Results on Financial PhraseBank Dataset

Model	Configuration	Metrics			
		Accuracy	*Precision*	*Recall*	*F1-score*
GPT-3.5-turbo	Zero-shot	0.78	0.79	0.84	0.80
	Few-shot	0.77	0.78	0.84	0.79
GPT-4o	Zero-shot	0.85	0.83	0.86	0.84
	Few-shot	0.86	**0.86**	0.84	0.85
FinBERT	Fine-tuned	**0.88**	0.85	**0.89**	**0.87**

Source: Shen, Zhang (2024)
arXiv:2410.01987

Text Generation & Summarization

Pre-LLM Era

- **Text Summarization**
  - Extractive Summarization
  - Selecting key sentences directly from the text.
  - Limited to surface-level understanding.
  - Disjointed summaries with minimal contextual coherence.
- **Text Generation**
  - Template-Based or Rule-Based Generation
  - Lacked creativity and adaptability
  - rigid outputs unsuitable for nuanced topics.

After LLMs

- **Text Summarization**
  - *Abstractive Summarization*
  - Text that conveys the core meaning
  - Generate context-aware summaries.
- **Text Generation**
  - *Dynamic and Contextual Generation*
  - Produces creative, adaptable outputs
  - Suitable for complex or sensitive topics (e.g., finance, healthcare).

LLM Context Window

Massive Context Window

- Traditional NLP: ~512 tokens
- LLMs like GPT-4: up to **32,000 tokens**
- Allows analysis of entire documents (e.g., earnings call + financials)

Human-Like Comprehension

- Connects information across sections, detects tone shifts
- Emulates the nuanced approach of a seasoned financial analyst

Redefining Financial Analysis

- Holistic Document Understanding
- Retains context throughout earnings calls, linking themes from opening remarks to Q&A

Decision Framework for NLP Solutions

When to Choose Classical NLP vs. LLMs

Classical NLP is ideal when

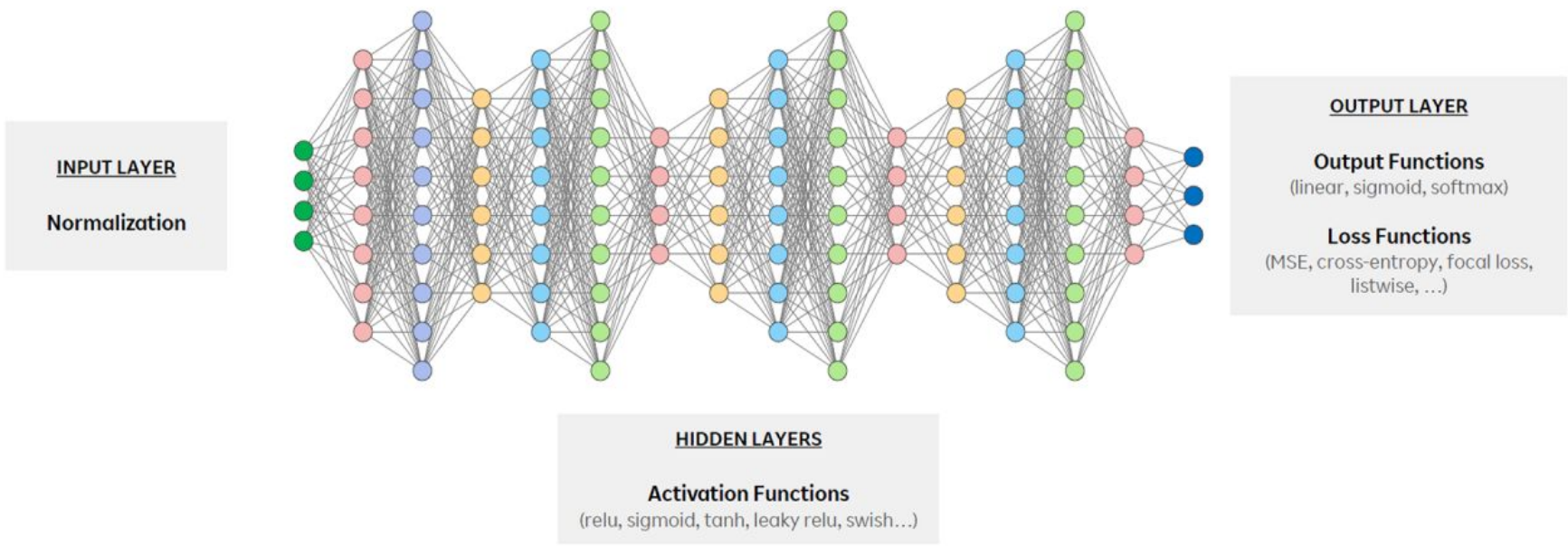
- Task is well-defined
- consistent outputs are required
- Resources are limited (low cost, fast inference, low memory)
- Transparency & explainability are essential
- Real-time processing is needed

LLMs are ideal when

- Complex understanding or generative tasks are required
- Flexibility across varied inputs is needed
- Fast development with minimal initial setup is beneficial
- Nuanced, creative tasks or handling edge cases are a priority

(Hybrid approaches can combine these strengths to maximize effectiveness)

Thank You!


INPUT LAYER
Normalization
OUTPUT LAYER
Output Functions
(linear, sigmoid, softmax)
Loss Functions
(MSE, cross-entropy, focal loss, listwise, ...)
HIDDEN LAYERS
Activation Functions
(relu, sigmoid, tanh, leaky relu, swish...)