RE-Work: AI in Finance Summit 2025

Natural Language Processing(NLP) in Fintech

Leonard Hawkes
Software Engineer, Inventor, Public Speaker

April 16, 2025

Agenda

I’m Leo - a Frontend Software Engineer. I’m passionate about new technologies, and following the latest trends in tech. Outside of code, you can catch me spinning tracks as my alias DJ Doc Aux all across the country.

- **Unstructured Data Decoding**
- **LLM-Powered Financial Products**
- **NLP Trends & Open-Source Tools**
- **Ethical AI & Regulatory Challenges**
- **Future of Autonomous Finance**
- **Q&A + Sources**

What is NLP?

Natural Language Processing (NLP) is a subfield of computer science and Artificial Intelligence (AI) that uses machine learning to help computers understand and communicate in human language.

Examples of NLP in Fintech	
BloombergGPT	Upstart
Erica (Bank of America)	AlphaSense
Cleo	Tractable

Unstructured Data Decoding

What is Unstructured Data?

Data that lacks a predefined format, and can't be easily stored in databases

Over 80% of financial data exists in unstructured formats including emails, reports, social media, and news articles.

This data is rich in insights but difficult to analyze without advanced technologies like NLP

Unstructured Data Decoding

The Unstructured Data Problem vs Solution

Over 80% of financial data exists in unstructured formats including emails, reports, social media, and news articles.

This data is rich in insights but difficult to analyze without advanced technologies like NLP

Text mining, or text data mining, is a data science discipline that employs natural language processing (NLP), artificial intelligence (AI), and machine learning to extract qualitative information from unstructured text.

Text analysis enhances this by identifying patterns in large datasets, yielding quantitative results.

Unstructured Data Decoding

Unstructured Data Use Cases

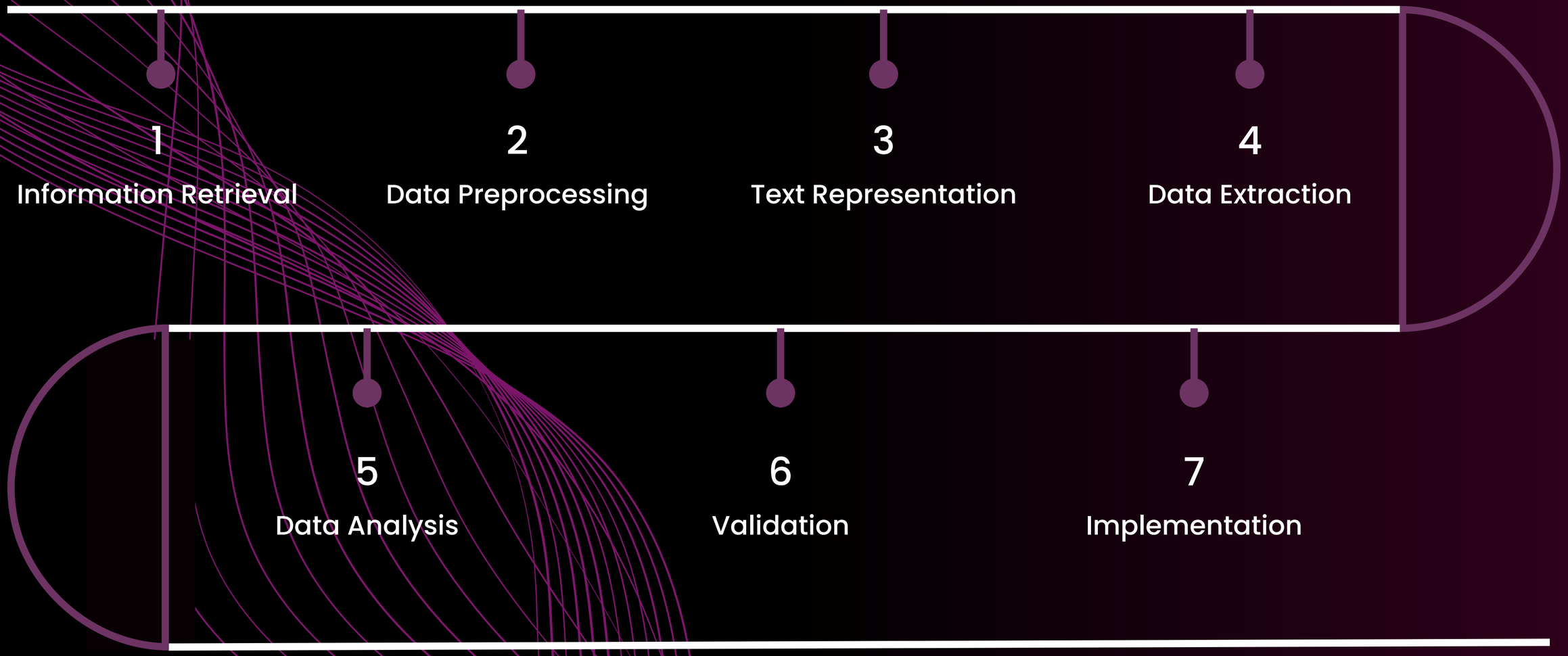
Risk Management: LLMs flag anomalies in loan applications
Market Predictions: Sentiment analysis of earnings calls predicts stock swings

Unstructured Data Decoding

Structured vs Unstructured Data

	Structured Data	Unstructured Data
Format	Structured data has a fixed format	Unstructured data does not
Storage	Structured data uses relational databases	Unstructured data requires specialized systems
Use Cases	Structured data is suitable for machine learning	Unstructured data is better for NLP and generative AI
Complexity	Structured data is easier to analyze	Unstructured data needs specialized tools

Text Mining Workflow



NLP in Fintech:
The Four Superpowers

01

Intelligent Data Extraction & Summarization

02

Sentiment Analysis for Market Predictions

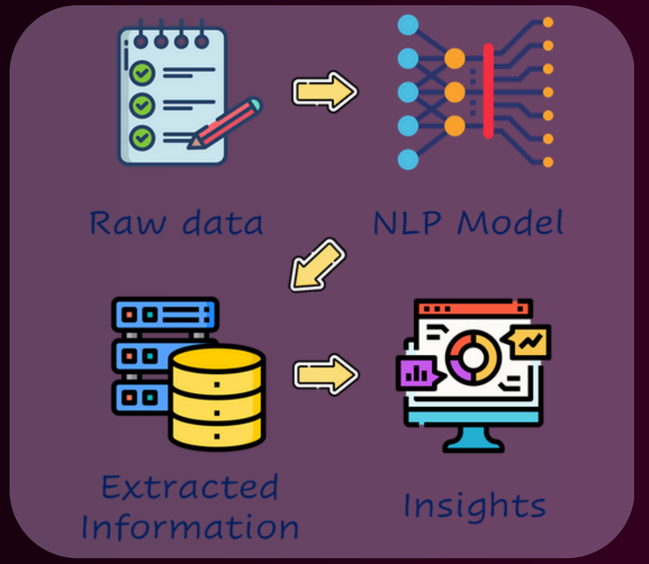
03

Automation of Customer Support

04

Risk Detection & Fraud Prevention

Intelligent Data Extraction & Summarization



What It Solves

Manual analysis of dense financial documents(10-Ks, contracts, earnings calls)

How LLMs Help

Extract key clauses, financial metrics, and obligations

Summarize 100-page reports into 3 bullet points

Examples

Some financial institutions review 12,000 anual contracts in seconds

AlphaSense: Uses NLP to highlight “supply chain risks” in SEC filings

Sentiment Analysis for Market Predictions

What It Solves

Human bias and latency in interpreting market-making news

How LLMs Help

Analyze earnings call transcripts, news and social media for bullish/bearish signals

Predict stock volatility with 85%+ accuracy (MIT Sloan)

Examples

BloombergGPT: Flags tone shifts in Fed statements for traders

Hedge Funds: Use NLP to parse Reddit/WSB sentiment for meme stocks

Automation of Customer Support & Financial Advice

What It Solves

High-cost, low-efficiency, human-led support

How LLMs Help

Chatbots handle FAQs, account queries, and personalized budgeting tips

AI advisors generate hyper-personalized investment memos

Examples

Erica (Bank of America): 50M+ interactions/year, 25% call center reduction

Morgan Stanley: AI drafts client reports from research notes

Risk Detection & Fraud Prevention

What It Solves

Rising financial fraud ($485B lost globally in 2023)

How LLMs Help

Detect phishing language in emails

Flag anomalous transactions (sudden large withdrawals)

Examples

Mastercard's AI: Reduces false declines by 30%

Darktrace: NLP analyzes trader communications for insider threats

Tools & Ethics: Open-Source Stack

FinBERT: NLP for Financial Sentiment Analysis

```
async function query(data) {
    const response = await fetch(
        "https://router.huggingface.co/hf-inference/models/ProsusAI/finbert",
        {
            headers: {
                Authorization: "Leonard_Hawkes",
                "Content-Type": "application/json",
            },
            method: "POST",
            body: JSON.stringify(data),
        }
    );
    const result = await response.json();
    return result;
}

query({ inputs: "This is such a cool presentation!" }).then((response) => {
    console.log(JSON.stringify(response));
});
```

FinBERT is a pre-trained NLP model designed specifically for financial sentiment analysis. It helps process financial documents, news, and reports to extract market sentiment

- Developed by ProsusAI and trained on financial texts.
- Helps analyze earnings calls, news, and reports for sentiment classification (positive, neutral, negative).
- Used by hedge funds, banks, and fintech firms for market prediction and investment insights.
- Used to analyze financial documents and predict market sentiment

Tools & Ethics: Bias & Fairness

When AI Reinforces Inequality in Finance

Equity in AI

Mortgage AI models required Black or Hispanic applicants to have credit score ~120 points higher than white applicants for equal approval chances*

Gender Inequality

Gender bias in credit scoring models: Historical data leads to women being offered smaller credit limits than men with similar profiles.

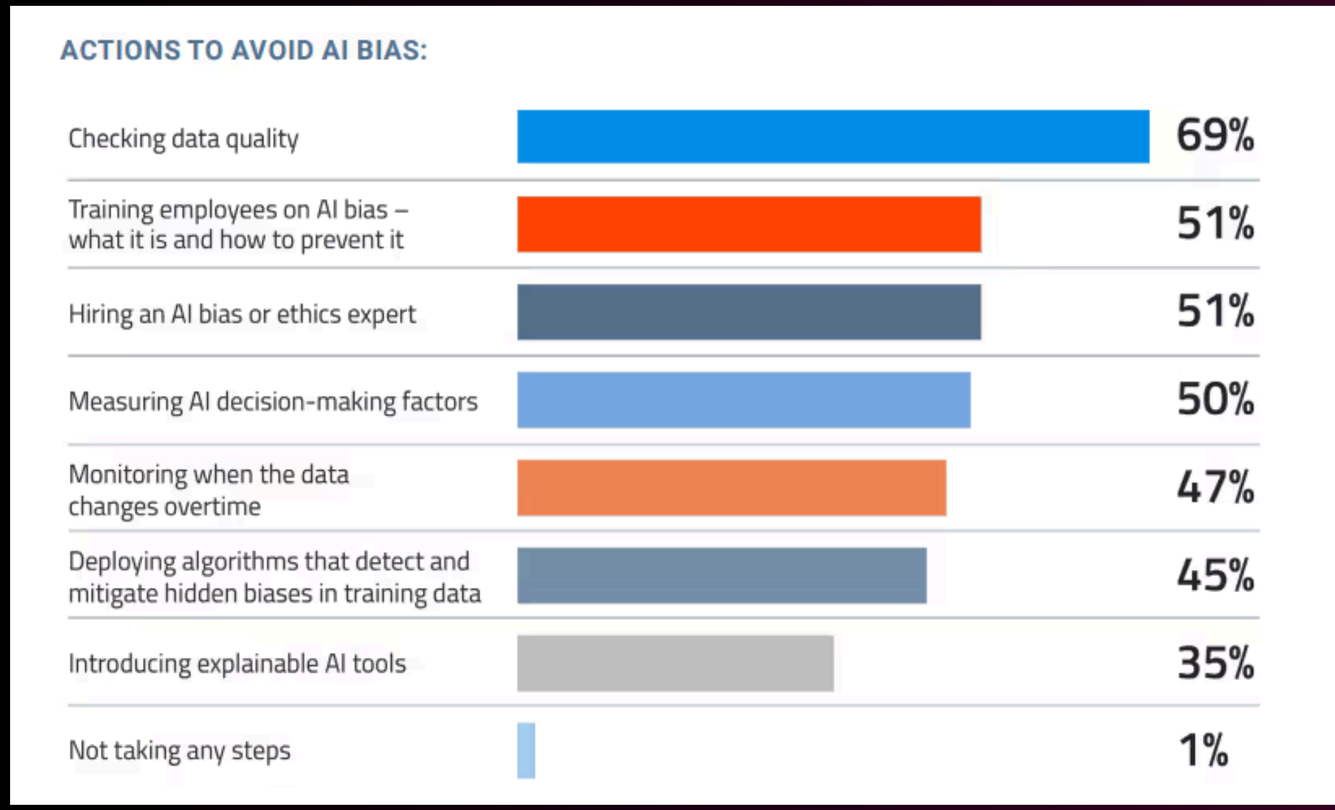
Financial Consequences

$2.2M settlement over AI screening software that disproportionately rejected Black rental applicants

Tools & Ethics: Bias & Fairness

The trade-off between predictive accuracy and AI decision transparency is a key challenge.

Complex models may yield higher accuracy but often act as "black boxes," making their decision-making processes unclear to users.



Tools & Ethics: Regulatory Challenges

Integrating LLMs like GPT-4 in financial services poses regulatory challenges, especially in the EU.

EU AI Act and Financial Services

regulates AI by categorizing systems based on risk levels. AI used in financial services, especially for credit assessments, is considered high-risk and must adhere to strict requirements for risk management, transparency, and accountability.

SEC Compliance and Regulation B

In the U.S., the Equal Credit Opportunity Act (ECOA) and Regulation B require lenders to give specific reasons for adverse actions like loan denials, promoting transparency and preventing discrimination in credit decisions.

Tools & Ethics: Regulatory Challenges

Can LLMs Explain Loan Denials Under Regulation B?

LLMs have the potential to assist in generating explanations for loan denials by analyzing application data and articulating reasons aligned with Regulation B requirements. However, several challenges must be addressed:

Accuracy & Fairness	LLMs must be trained on unbiased data to prevent perpetuating existing biases
Transparency	The decision-making process of LLMs should be interpretable, allowing clear and understandable explanations to applicants
Regulatory Alignment	Outputs from LLMs must comply with legal standards, providing specific and actionable reasons for credit decisions

Tools & Ethics: Cost vs Value

LLMs in Fintech: Is the Value Worth the Cost?

💸 The Cost:

Training & Fine-Tuning: Millions of dollars (e.g., GPT-4 cost over $100M).

Infrastructure: Requires high compute power (GPUs), storage, and low-latency architecture.

Inference Costs: Serving LLMs via APIs can be costly, especially for real-time applications.

Compliance & Audit Overhead: Implementing safeguards increases time and expenses.

📈 The Value:

Automated Insights: Quickly analyze sentiment, intent, and trends from financial data.

Enhanced Customer Support: Chatbots lower costs and increase satisfaction, exemplified by Klarna's reduced service tickets.

Regulatory Assistance: Help generate explanations for regulatory decisions.

Product Innovation: Facilitate new offerings like personalized investment summaries and fraud alerts.

Conclusion

Natural Language Processing is no longer just a backend tool — it's becoming the voice of modern finance.

Sources

Google Doc detailing some of my sources

Questions & Discussion

I'd love to hear your thoughts, experiences, or questions about how LLMs are transforming fintech — especially in your own work