



## 1. Data Collection & Preprocessing Tools

AI services begin with large-scale data extraction from structured and unstructured sources:

- **Web Scraping Tools:** Selenium, iTEXT-PDF  
→ Used to collect financial news, analyst reports, and company announcements.
- **PDF/Text Parsers:** iTEXT-PDF Adobe PDF Extract  
→ Used to extract tables and text from annual reports, audit documents, etc.
- **OCR Tools:** Tesseract  
→ For scanned documents like signed financial statements.
- **ETL Pipelines:** Jenkins  
→ Used for ingestion, transformation, and loading of diverse data types.

## 2. Natural Language Processing (NLP) Algorithms

Used for analyzing unstructured data like reports, earnings call transcripts, and data.

- **Named Entity Recognition (NER):**  
→ To detect company names, financial events, management changes, etc.
- **Sentiment Analysis:**  
→ Proprietary logic for identifying positive/negative outlooks in news or earnings calls.
- **Text Summarization:** iTEXT-PDF  
→ To create executive summaries of long reports.

## 4. Machine Learning Models for Scoring & Classification

Used to assess company health or categorize risk/performance.

- **Random Forests / Gradient Boosted Trees : LightGBM:**  
→ Used to build models predicting financial distress, stock rating, or scores.
- **Clustering (K-Means, DBSCAN):**  
→ For grouping companies with similar risk or performance patterns.
- **Anomaly Detection Models:** Isolation Forests, One-Class SVM  
→ To flag sudden changes in balance sheets or market behavior.

\*\*\*\*\*