

COURSE 1 PROJECT

Project Title: J P Morgan classification for legal documents

Phase 1: Business Understanding

- 1. Defining business objective:** J P Morgan developed a new software called **COIN**. The objective of the COIN is to enhance document review and reduce manual efforts .It also reduce the time and cost associated with manual document review and increase the efficiency of legal classification .
- 2. Business goal:**Automate the classification of various legal documents.

Phase 2: Data Understanding

The data used in COIN for automation comes from loan agreements , legal contract databases, past legal agreements . These data may also contain clauses , parties involved and contract type.

Phase 3: Data Preparation

The data preparation steps involve removing unwanted data (data cleaning) , converting all the documents to machine-readable text, and labelling to assign predefined categories to contracts based on legal definitions.

Phase 4: Modeling

We have many machine learning approaches which can be used(both unsupervised & supervised). Firstly , we will be grouping the data and follow topic modelling for automating the categories . This process involves splitting data as Training & Testing .

Phase 5: Evaluation

In this step we will evaluate the result(automated classification result) and compare the result with manual classification accuracy which helps in assessing model effectiveness . If the result is not effective then the modeling will be done again. Then error analysis should be also done in this phase.

Phase 6: Deployment

The model can be integrated as a cloud-based API for legal teams, connected with document management systems, and presented through an interactive dashboard for real-time monitoring. To ensure long-term efficiency, continuous model training with new data, regular audits, and compliance checks are necessary, along with updating legal contract categories based on regulatory changes.