

Business Understanding

Problem statement - Deploying COIN software to review more complex legal document

Goals – By using CRISP -DM method train software for review complex legal document

Data Understanding

- First, I will gather all the initial data which is needed to fulfill business objectives
- This data gathering includes a more complex bank legal document and specific document type that is needed to fulfill business objectives like credit-default swaps, custody agreements, etc.
- Get data of approve document and disapprove document and their related files and data
- I will maintain different types of data files for specific legal documents.
- Ensure quality of source and quality of data

Data Preparation

- Convert all unstructured data to structured and systematic
- Arrange data in different columns as per different content of document
- Remove duplicate in unique data, remove any unwanted gaps, space and format it neat and clean tables
- Remove error in data

Modeling

- Here involve training data set so it will work for more complex legal documents
- Using various tools to find patterns in a data set for approving and disapproving particular documents
- Running algorithm and machine learning to give instruction and training to software for different possibility of approving and disapproving document

Evaluation

- After training software testing its reliability by considering different aspects of legal documentation
- Ensuring our model works well on different clauses of variety of legal documents by converting different clauses to mathematical expressions and then testing

Deployment

- After Evaluation Deploying our Updated software to real word scenario
- After running model in business verifying is the model completing business objective and goals of our data mining
- If it is not solving business problems again go back and debug in the cyclic process of CRISP-DM method.