

PROJECT: - J P Morgan classification for legal documents

Problem Statement: - How would you break the problem/scenario down with the help of CRISP-DM methodology: "Automate the classification of various legal documents".

Solution: - by using CRISP-DM methodology breaks the business problem into steps given below

1 Business understanding: - so JP Morgan company work on various legal documents for their lawyers the basic objective of this project is to classified that document in their specific categories. As it should be considering current situation and business model of JP Morgan firm required to provided all data for further proceeding.

2 Data understanding: - Initially we required all data from firm it includes all legal documents. So, all legal documents should be in scanned copy that is image resolution OCR form. We have to collect all data form JP Morgan firm. Data is in Unstructured form. The source of this data is machine data source because its scanned copy of documents and it is in image form so required big storage. We have to collect image form of all legal documents and stored it in databased.

3 Data preparation: - This phase is dedicated to cleaning and transforming raw data into a suitable data so those raw image form data we need to clean as per business requirements. Need to select an image based on various category of legal document and arranged it in proper manner after that we need to delete or clear those image that are repeatedly occur based on pattern of image such as any word or location in the image. Present all image in various categories of document in databased form.

4 Modeling: - with clean data in hand, various modeling techniques are applied here based on specific data format. So here as data is in image form, we have to apply some modeling techniques. So, selecting some modeling techniques then design a test to run this model as per business requirement.

5 Evaluation: - After completing all requirement and formatting data in proper way and designing a suitable model to run this project we have to check weather this fulfilled the requirement and result of this process should matched with businesses' objective

6 Deployment: - this is final step of this process after monitoring and maintenance reviewing the result of the project finalized the project to business to run this model.

Video Link: - <https://drive.google.com/file/d/10jU5OTsfgrx86zHEE054J55GHc0T6R-/view?usp=drivesdk>