

Project work

DISTINCT THE PROCESS BASED ON ALL THE STEPS IN CRISP-DM:

CRISP-DM:

CRISP-DM, or Cross-Industry Standard Process for Data Mining, is a cyclical framework that offers a systematic approach for planning, organizing, and executing a data mining project.

Business Understanding:

Objective:

JP Morgan aims to automate document reviews to save time, enhance accuracy, and reduce errors.

Current Situation:

The legal and banking offices are experiencing delays and potential mistakes due to manual processes.

Goal:

The goal is to complete document reviews in seconds, improving accuracy while minimizing errors by automating the review process.

Project Plan:

A structured plan to implement automated document reviews for a specific class of contracts.

Data Understanding:

Gathering Data:

Collect data relevant to automated documents, such as legal documents, commercial agreements, and custody agreements.

Describing Data:

Determine which data is essential for the automated document process in JP Morgan's project.

Exploring Data:

Analyze the structure of the documents and identify recurring patterns within the text.

Verifying Data Quality:

Assess the quality of the data and patterns to ensure they meet the requirements of JP Morgan's project.

Data Preparation:

Selecting Data:

Choose the data relevant to the JP Morgan project.

Select Documents and Services:

Identify the appropriate documents and services for JP Morgan.

Cleaning Data:

Remove irrelevant or duplicate data. Extract meaningful information from JP Morgan's provided data.

Integrating Data:

Combine data from different sources into a single source, categorizing and grouping documents like custody agreements, legal documents, etc.

Formatting Data:

Organize and arrange the data according to the project's requirements.

Modelling:**Selecting Modelling Techniques:**

Choose the appropriate model for training data related to the automated document review for the JP Morgan project.

Designing and Testing:

Design the model using algorithms to address the automated review document requirements for the JP Morgan project. Test the model to evaluate its accuracy and performance in data assessment.

Building the Model:

Train the model using labeled data, such as documents and agreements, to teach it how to recognize patterns.

Accessing the Model:

Enable real-time usage of the model for automated legal document processing, improving process efficiency.

Evaluation Process:

1. **Evaluation:** Assessing the results of the model used by JP Morgan to automate the classification of legal documents.
2. **Reviewing the Process:** Monitoring the model's performance in automating the document review process and evaluating its functionality.
3. **Determining the Next Steps:** After completing the review, deciding on any necessary model adjustments to ensure continuous improvement.

Deployment:

1. **Planning Deployment:** Developing a plan to implement the model in a real-time environment for the JP Morgan project.
2. **Monitoring and Maintenance:** Tracking the model's performance in the real-time environment and ensuring it maintains accuracy.
3. **Review of Project:** Continuously assessing the project's progress in automating the classification of legal documents.
4. **Finalizing the Project:** If the project review is successful, proceeding with the final launch to fully automate the classification of legal documents.