AI Project stages

Based on your learning of different stages of Ai Project. In this assignment you are supposed to write an article describing different stages of an Ai-Project and role of an Ai-Engineer/Researcher in the same.

1. You can refer the sentiment analysis slide in Data folder for more help.
2. For explaining the different stages using an example feel free to choose a use case of your choice/interest.

Example:

O [OSEMN]: Obtain

1. Description:
2. Understanding business problem is crucial for any projects progress. In “O” stage of OSEMN process engineers are responsible for Understanding the Business Problem and Collecting the Relevant Data.
3. Business Problem understanding can include following aspects:
4. Types of Problem
5. Outcome
6. Collecting Data can include following activities:
7. Identify the source of data
8. Specifying relevant data requirement. 1. Format, 2. Data Size.
9. Role of an AI – Engineer:
10. Not essentially in business negotiation but an Ai-Engineers role is crucial here to understand the problem statement, scoping the project, finalizing the timelines, informing about the challenges to business heads.
11. An Ai-Engineer is not responsible for collecting the data but might get involved in specifying the data requirement and verification of data.

# **Stages of an AI Project and the Role of an AI Engineer/Researcher**

Developing an AI project involves a structured, methodical approach, typically broken down into stages. These stages ensure that the AI solution is effective, reliable, and aligned with the business goals. The OSEMN framework—Obtain, Scrub, Explore, Model, and Interpret—provides a comprehensive blueprint for these stages. Using a sentiment analysis project for customer reviews as an example, and describe the role of an AI Engineer or Researcher in each stage.

# **Obtain (O): Understanding the Business Problem and Data Collection**

The first step in any AI project is understanding the business problem. For our example, let’s say a business wants to analyse customer feedback to improve product satisfaction. The goal is to create a sentiment analysis model to classify customer reviews as positive, negative, or neutral.

## **Understanding the Business Problem**

AI Engineers work closely with business stakeholders to understand:

* The type of problem (classification, prediction, etc.)
* Desired outcomes (e.g., sentiment scores or categorical labels for reviews)

## **Collecting Relevant Data**

Once the problem is defined, the AI Engineer or team:

* Identifies data sources (e.g., review databases, customer feedback forms)
* Specifies data requirements, including format (text data) and size, to ensure enough data for accurate model training

## **Role of an AI Engineer**

Although not usually involved in business negotiations, the AI Engineer plays a crucial role in:

* Understanding the project scope, estimating timelines, and identifying potential challenges.
* Collaborating on data requirements with data collection teams to ensure that the dataset meets quality standards for the project.

# **Scrub (S): Data Cleaning and Preparation**

Data scrubbing or cleaning is critical in ensuring that the model learns from accurate, well-structured data. In sentiment analysis, raw data often includes irrelevant information, typos, and inconsistencies, which can impact model accuracy if left unaddressed.

## **Cleaning Text Data**

Text data needs extensive pre-processing, including:

* Removing unnecessary characters, correcting misspellings, and handling emojis or special characters.
* Normalizing text by converting it to lowercase, removing stopwords (e.g., "and," "the"), and tokenizing it.

## **Handling Missing Values**

Missing or irrelevant data points are managed through:

* Removal or imputation of missing values.
* Identifying noise (like non-informative reviews) and either cleaning or removing it.

## **Role of an AI Engineer**

The AI Engineer plays a hands-on role in:

* Pre-processing and cleaning the text data, ensuring uniformity and accuracy.
* Implementing scripts to automate cleaning tasks and validating that data quality aligns with project goals.

# **Explore (E): Data Exploration and Analysis**

Data exploration helps uncover patterns, correlations, and key characteristics of the dataset. This step is essential for understanding customer sentiment distributions and common expressions in positive vs. negative reviews.

## **Statistical Analysis and Visualization**

* Performing exploratory data analysis (EDA) to identify common words, average review length, and sentiment trends.
* Visualizing data with word clouds, bar charts, or sentiment distribution histograms to provide insights into customer feedback.

## **Feature Engineering**

* Creating new features (e.g., sentiment polarity scores, review length) that could enhance model performance.
* Experimenting with representations like TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings to capture meaningful aspects of the text data.

## **Role of an AI Engineer**

* The AI Engineer actively conducts EDA, identifies potential features, and interprets patterns that may influence model design.
* This role includes communicating key findings back to the business team, providing context on how certain sentiment characteristics might relate to the customer experience.

# **Model (M): Model Building and Evaluation**

After the data is prepared and explored, the model-building phase begins. This phase involves selecting, training, and fine-tuning machine learning or deep learning models to best classify sentiment in customer reviews.

## **Model Selection and Training**

* Selecting suitable models (e.g., Naïve Bayes, Support Vector Machines, or deep learning models like LSTM or BERT).
* Training the models on pre-processed data, optimizing for accuracy, precision, and recall.

## **Model Evaluation**

* Evaluating the model's performance using metrics such as accuracy, F1-score, and confusion matrix.
* Conducting cross-validation to ensure the model’s generalizability.

## **Role of an AI Engineer**

* The AI Engineer is primarily responsible for selecting appropriate models and implementing the training process.
* They experiment with hyperparameters, apply regularization techniques, and iterate on models to enhance performance.
* The AI Engineer also validates the model’s output, making adjustments as necessary to align with business requirements.

# **Interpret (N): Model Interpretation and Business Integration**

In the final stage, the model’s results are interpreted, and the insights are shared with stakeholders. The focus is on understanding the model’s predictions and integrating its outputs into the business context for practical use.

## **Model Interpretation**

* Interpreting model predictions by analysing sentiment breakdowns, key drivers of negative feedback, and topics associated with customer satisfaction.
* Using interpretable models or techniques (e.g., SHAP, LIME) to explain predictions.

## **Business Integration and Feedback**

* Deploying the model into the production system, allowing real-time or periodic sentiment analysis.
* Gathering feedback from end-users and business teams to improve model performance iteratively.

## **Role of an AI Engineer**

* The AI Engineer ensures the model’s outputs are understandable and valuable for decision-making.
* They oversee deployment, monitor model performance in production, and handle any required maintenance, making adjustments based on feedback.
* The Engineer also assists in creating visualizations or dashboards to allow stakeholders to interact with the insights generated.

## **Summary of Roles Across Stages:**

|  |  |  |
| --- | --- | --- |
| Stage | Key Activities | AI Engineer's Role |
| Obtain | Understand the problem, collect relevant data | Scope project, define data requirements, and work with data teams |
| Scrub | Clean and pre-process data | Execute data cleaning, create pre-processing pipelines |
| Explore | Analyze data and identify features | Conduct EDA, feature engineering, and pattern identification |
| Model | Train and evaluate models | Choose models, tune parameters, evaluate accuracy, iterate |
| Interpret | Interpret results, deploy, and integrate with business processes | Ensure model interpretability, oversee deployment, gather feedback for improvement |

This structured approach and the active involvement of an AI Engineer or Researcher are crucial in guiding an AI project from concept to deployment, ensuring that each stage is aligned with the project’s overall objectives and delivers value to stakeholders.