CHATBOT IN PYTHON

Phase 4 : Development Part 2

Feature Engineering :

Feature Engineering is the process of creating new features or transforming existing features to improve the performance of a machine-learning model. It involves selecting relevant information from raw data and transforming it into a format that can be easily understood by a model. The goal is to improve model accuracy by providing more meaningful and relevant information.

Feature engineering is the process of **transforming raw data into features that are suitable for machine learning models**. In other words, it is the process of selecting, extracting, and transforming the most relevant features from the available data to build more accurate and efficient machine learning models.

The success of machine learning models heavily depends on the quality of the features used to train them. Feature engineering involves a set of techniques that enable us to create new features by combining or transforming the existing ones. These techniques help to highlight the most important patterns and relationships in the data, which in turn helps the machine learning model to learn from the data more effectively.

Model Training :

Model training is the phase in the data science development lifecycle where practitioners try to fit the best combination of weights and bias to a machine learning algorithm to minimize a loss function over the prediction range.

The purpose of model training is to build the best mathematical representation of the relationship between data features and a target label (in supervised learning) or among the features themselves (unsupervised learning). Loss functions are a critical aspect of model training since they define how to optimize the machine learning algorithms. Depending on the objective, type of data and algorithm, data science practitioner use different type of loss functions. One of the popular examples of loss functions is Mean Square Error (MSE).

C3 AI enables distributed training through a mix of out-of-the-box and custom ML pipelines addressing different data science workload demands. The training of these pipelines creates ML models which can be analyzed in the C3 AI ML Studio, promoted for deployment, used for generating score reports, or evaluating model performance. Further these models could also be created using no-code drag-and-drop experiences provided by C3 AI Ex Machina.

Evaluation :

Evaluation is the structured interpretation and giving of meaning to predicted or actual impacts of proposals or results. It looks at original objectives, and at what is either predicted or what was accomplished and how it was accomplished. So evaluation can be formative, that is taking place during the development of a concept or proposal, project or organization, with the intention of improving the value or effectiveness of the proposal.

Evaluation is inherently a theoretically informed approach (whether explicitly or not), and consequently any particular definition of evaluation would have been tailored to its context – the theory, needs, purpose, and methodology of the evaluation process itself. Having said this, evaluation has been defined as:

- A systematic, rigorous, and meticulous application of scientific methods to assess the design, implementation, improvement, or outcomes of a program. It is a resource-intensive process, frequently requiring resources, such as, evaluate expertise, labor, time, and a sizable budget[4]
- "The critical assessment, in as objective a manner as possible, of the degree to which a service or its component parts fulfills stated goals" (St Leger and Wordsworth-Bell).[5][*failed verification*] The focus of this definition is on attaining objective knowledge, and scientifically or quantitatively measuring predetermined and external concepts.
- "A study designed to assist some audience to assess an object's merit and worth" (Stufflebeam).[5][*failed verification*] In this definition the focus is on facts as well as value laden judgments of the programs outcomes and worth.