To,

IITD-AIA Foundation of Smart Manufacturing

<center>Subject: **Weekly Progress Report for Week 1**</center>

Dear sir,

Following is the required progress report to the best of my knowledge considering relevant topics to be covered.

What's happening this week:

- Overfitting and Underfitting
- Dropout and Batch normalization
- Different kinds of chatbots
- AI enabled chatbots
- Python libraries like NLTK
- Generative modelling
- Stochastic Gradient Descent

My Understanding of INTP23-ML-01: Chatbot for FSM

Scope:

This project involves developing an interactive conversational interface that can assist users by providing instant responses to their queries, guiding them through the website, and resolving common issues. The chatbot should be equipped with natural language processing capabilities, allowing it to understand and respond to user input accurately. It can offer student support, retrieve information from the website, provide student with the information he wants, guide him to the website of fsm.

Solution:

This kinda chatbot can be made using natural language processing (NLP) where it involves designing and implementing an intelligent conversational bot that can understand and respond to user queries in a natural and helpful manner.

Approach:

It involves collecting all the data from the scratch of the website and then designing the conversational flow and user interface. NLP and MI models will be applied to make the bot work with different languages.

s

Weekly Progress:

June 05:

Covered some topics like overfitting and underfitting with hands on session on Kaggle.

- Overfitting is a common problem in DL where a model performs exceptionally well on the training data but fails to generalize well to new, unseen data and occurs when the model learns to memorize the training examples instead of learning the underlying patterns and relationships.
- Underfitting is the opposite problem to overfitting and occurs when the model fails to capture the underlying patterns and relationships in the training data, resulting in poor performance on both the training data and new, unseen data
- Underfitting the training set is when the loss is not as low as it could be because the model hasn't learned enough signal. Overfitting the training set is when the loss is not as low as it could be because the model learned too much noise.
- Graph of overfitting and underfitting
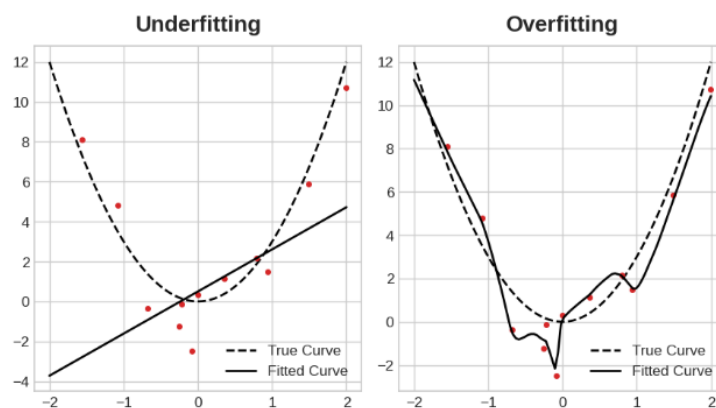
June 06:

Learnt Techniques that prevent overfitting and underfitting in the training of a model.

- Dropout and batch normalization are the techniques used to improve model performance, prevent overfitting, and enhance training stability.
- Dropout is a regularization technique that aims to reduce overfitting by randomly dropping out (zero) a proportion of neurons in a neural network during training. The idea behind dropout is to break up these weights, we randomly drop out some fraction of a layer's input units every step of training, making it much harder for the network to learn those spurious patterns in the training data.
- Batch normalization can help correcting training that is slow or unstable. It normalizes the inputs of each layer by subtracting the batch mean and dividing by the batch standard deviation. By normalizing the inputs, batch normalization reduces the internal covariate shift, which is the change in the distribution of network activations during training.
- Binary classification is a type of supervised learning task in machine learning where the goal is to predict a binary outcome or assigning a binary label to each input instance based on its features or characteristics.

June 07:

INTP23 allocation: Chatbot for FSM

- Discovered some resources for the project and studied the types of chatbots
- Namely, Rule based chatbots, retrieval based chatbots, generative chatbots, AI enabled chatbot, virtual assistant and many more.
- Learned Generative chatbot and AI enabled chatbot thoroughly.

June 08:

Now after deciding the type of chatbot, proceeding towards the NLP and Rasa.

- Natural Language Processing is the area of AI that creates a software which understands written and spoken language by humans and acts as an interpreter between computer and human language.
- NLP has a various use cases like text classification, intent classification, sentiment analysis, text summarization, Keyword Extraction, question answering and many more.
- Rasa Framework is a tool to build custom AI chatbots using Python and natural language understanding (NLU). It has Main 2 component i.e., Rasa NLU and Rasa Core.

June 09:

Studied python libraries: SciPy, PyTorch, Transformer, NLTK.

- NLTK (Natural Language Toolkit) is a Python library which is used for natural language processing (NLP) tasks and has resources to work with human language data, making it easier to process, analyse, and manipulate text.
- SciPy is also a open-source Python library used for scientific and technical computing. It is used for numerical operations, linear algebra, integration and various mathematical algorithms.
- PyTorch is a machine learning framework for Python that enables developers to build and train deep neural networks. It also fastens the process between research prototyping and deployment. Difference between PyTorch and TensorFlow is PyTorch creates dynamic graphs while TensorFlow creates static graphs.
- Created a flow diagram for upcoming tasks for making a chatbot.

June 10:

Learnt about different kinds of models:

- Generative modelling is the use of artificial intelligence, statistics and probability in applications to produce a representation of observed or target variables that can be calculated from observations. It is also used for unsupervised learning to predict all manner of probabilities on a subject from modelled data.
- 3 kind of generative modelling, namely: generative adversarial network (GAN), Variational Autoencoders(VAE) and Autoregressive models.
- GAN and VAE is based on neural network. VAE is the most efficient method for developing generative model as it seeks to understand the underlying probability distribution of the training data so that it can quickly sample new data from that distribution.

June 11:

Studied the gradient in deep learning with its hands on practice on Kaggle.

- Gradient descent is iterative algorithm which is used to minimize the loss function or any given function. There are 3 types of gradient descent namely, batch gradient descent, stochastic gradient descent and mini batch gradient descent.
- Gradient Descent is used for predicting discrepancy between models predicted value and true values for any trained model.
- Batch gradient descent is algorithm that uses entire training dataset at each iteration and update its parameter whereas Stochastic gradient descent starts by selecting few samples from the dataset instead of selecting the whole dataset for each iteration. and mini batch gradient descent is the combination above both gradients, it divides the dataset into small batch and computes the gradient.
- "Stochastic" means "random" so a process is linked with random probability.
- Working of the algorithm: Stochastic gradient descent (SGD) starts with computing the gradient function and picks a random initial value for the parameter, updates the gradient function for the parameter values. Then calculates the step sizes for each features and also calculate new parameters and repeat the process until gradient is almost 0.