**IITD Machine Learning-01(chatbot)[INTP23-ML-1]**

**Day 18:**

On day 18, of IITD-AIA FSM internship, I did scrape of IAFSM and stored in json file.

**Learning:**

Today, I first found pattern in which data is stored in “p” tags or directly written in “div” tag. Whenever content is directly written in “div” tag their class name is “sppb-addon-content”. So, I searched for this class named div and printed the file in json. Here, yesterdays’ error of repeating content is removed. But now error is, value is stored in different curly brackets. So, in order to remove error, I first used “read and write” method on file. However, it didn’t work as most of the data is removed using this method. Now, I have to change the code, where I have to write in json only once so in order to that I used dictionaries and list. Under I dictionary I used four lists, in which values like link, paragraph, heading, div content are appended and in the last, this dictionary is stored in json.

**Work Done/learning Implemented:**

I scraped out IAFSM website completely and stored data in json.

**Is Progress As per Track?**

Yes, the processing of the data, in order to apply NLP, EDA

**Issues Faced Today:** error in storing data in json

**Issues Closed Today:** putting the scrapped data into “json” file, without repetition, error in storing data in json

**Highlights:**

I stored the scraped data into a json file/ format of IAFSM website. However done the same with “FSM skills” website.

**Concluding the day:**

So today, I completed json file from scrapped data of IAFSM completely