

Task 3.2

REINFORCEMENT LEARNING

Reinforcement learning is a branch of machine learning along with Supervised and Unsupervised learning. It works on a very basic principle of feedback to a particular response (That is reward and the goal is always to gather more and more reward). feedback can also be negative which is called penalty. The machine learns from its mistakes and updates its system with the right answer so that whenever a similar question is asked to the machine, it responds in a correct manner.

Then a strategy is made to predict future actions by the experience that the machine has gained.

(went through the learning algo like Q learning and policy gradient but could understand very little)

For example in a very layman's language whenever a toddler touches a candle, he or she feels a burning sensation and remembers to not repeat the same mistake again. In a similar fashion, The machine identifies its mistakes and keeps on updating and learning from it therefore self-learning, improvisation and updating the system with the correct answer is a way in which reinforcement learning works.

Reinforcement learning is divided in 2 parts, Model based reinforcement learning and model free Reinforcement learning. In Simple words from what I understand, Model based Reinforcement learning is based on construction of a machine learning model which helps to predict the future whereas model free reinforcement learning relies on gaining experience and taking appropriate actions. (doubt)

There are a plethora of advantages to Reinforcement learning. For example:
It helps in continuous learning that means performance of the machine will increase because the machine is gaining experience overtime and helps to improve its performance to take actions. It is used to solve complex algorithms where manual behavior is not possible.

Applications of Reinforcement learning is in a lot of things:

1. While playing Fifa on the playstation when the user plays nicely, The machine or the computer alters the difficulty level on the level of play of the user. Here the machine learns from the experience by playing with the user and taking informed actions.
2. Reinforcement learning is also used in stock-trading. This algorithm is built on previous experiences of the data of the stock.

3. It is used in self-driving cars where machines learn by interacting with the environment.
4. It helps in recommendation systems on platforms like Netflix, Youtube, etc. Basically it gains experience what type of genre the user likes to watch and based on that, helps to recommend movies or videos.

Earlier chat gpt did not use reinforcement learning and only used supervised learning where they had a large set of labeled data and took decisions with respect to the given data. This was the main reason why it used to give bogus answers because they had no mechanism to correct the inaccuracy of the data.

(gpt2 and 3 were based on supervised learning)

(gpt 3 onwards, it is based on the concept of reinforcement learning)

Later on when Reinforcement learning was introduced, chatgpt became better as it could learn continuously and learn from its mistakes and store the correct data so that whenever a similar question is asked, it can give an accurate response.

I think chatgpt is still improving, it still has a lot of inaccuracies which might be rectified in the later version of gpt.

Task 3.1

Preprocessing of textual data is extremely necessary for normalizing the textual data so that it can be processed further and is devoid of any inaccuracies

It helps to maintain the quality of data and reduce any kind of noises or irrelevant information. Preprocessing ensures uniformity within the dataset, making it easier for algorithms to identify patterns and make accurate predictions.

The steps that will be followed for preprocessing will be:

- Lower casing- since system are case sensitive therefore all words are made to lower case to remove any kind of inaccuracy
- Removal of Punctuations since punctuations are bogus for the system and acts as noise for the system.
- Removal of Stopwords- words like article or pronoun are reduced so that accuracy of the system can be increased
- Lemmatization where unnecessary verbs of any word are reduced to its original form for better understanding of the machine. That is words are reduced to dictionary based simplest forms where words have a meaning unlike some cases of stemming. The base words which has been reduced is known as 'lemma'
- Stemming is done so that words are reduced to its base form. Although this process is fast, it can give answers that don't make sense according to the dictionary .

(Stemming and Lemmatization is widely used in NLP models)

- Removal of emojis is done to reduce any kind of unnecessary data which would not concern the machine and may increase inaccuracy if it is not relevant.
- Spelling correction is also done to reduce inaccuracies
- Removal of extra white spaces is also done to reduce the inaccuracy

These steps are extremely essential for textual preprocessing as it can reduce the inaccuracies present in the data thereby cleaning the data. It is also an essential step for NLP models as textual data has to be cleaned and pre-processed for analysis of the data so that it doesn't give rise to any kind of ambiguity.