# PROGRESS REPORT(PROJECT DIARY) :

Name: Ashish Gupta

Project: Unscripted

| DATE | PROGRESS |
|---|---|
| 01-04 April 2021 | <ul><li>Created a new GitHub account</li><li>Created a new Colab notebook</li></ul> |
| 05-11 April 2021 | <ul><li>Completed the following videos on Neural Network:<br>https://youtu.be/aircAruvnKk<br><br>*Introduction to the Neural Networks. Learned about the structure of artificial neural networks, different layers of neural networks, and how each layer is connected by means of weights and biases and an activation function finally.*<br><br>https://www.youtube.com/watch?v=IHZwWFHWa-w<br>*Learned about the cost function, backpropagation and gradient descent. The cost function is the function that tells how far the network's prediction is from the actual target value, and backpropagation is the algorithm for adjusting the weights to reach to target prediction by calculating the gradients. And gradient descent is the way to converge towards some local minima of the cost function.*<br><br>https://www.youtube.com/watch?v=Ilg3gGewQ5U<br>*Learned about what backpropagation actually does. Different gradient descent techniques are- Batch Gradient Descent, Stochastic Gradient Descent (SGD), and Mini-batch Gradient Descent.*</li><li>Completed videos covering Numpy, Pandas and Matplotlib from the following playlist:<br>https://www.youtube.com/watch?v=bPrmA1SEN2k&list=PLZoTAELRMXVPBTrWtJkn3wWQxZkmTXGwe&index=1</li><li>Covered the first 2 chapters of the following book:<br>https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/<br>*The first chapter of this book gives the introduction of Machine Learning, Different systems of Machine Learning, and Challenges in ML.*<br>*The second chapter talks about implementing an end to*</li></ul> |

| | |
|---|---|
| | *end regression task in ML which includes visualizing the data and cleaning it accordingly and preparing for building the model.  Setting up an environment in Jupyter Notebook, splitting data into train and test data, building the model and training and testing it.* |
| 10-16 May 2021 | ● Covered the third and fourth chapters of the following book: https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/ *In these chapters, I learned about the classification in ML using MNIST dataset, and gradient descent techniques.* |
| 17-23 May 2021 | ● Completed first two videos of course CS224N: NLP with Deep Learning https://youtube.com/playlist?list=PLoROMvodv4rOhcuXMZkNm7j3fVwBBY42z |
| 24 May - 04 June 2021 | ● Was having seminar work from 16th May to 04th June. |
| 05-13 June 2021 | ● Read the following suggested blogs on Automatic Speech Recognition: https://heartbeat.fritz.ai/a-2019-guide-for-automatic-speech-recognition-f1e1129a141c https://www.analyticsvidhya.com/blog/2021/01/introduction-to-automatic-speech-recognition-and-natural-language-processing/ https://medium.com/georgian-impact-blog/how-to-make-an-end-to-end-automatic-speech-recognition-system-with-wav2vec-2-0-dca6f8759920 <br><br>● Implemented Watson Speech recognition package from IBM on the given audio file. |
| 14-20 June 2021 | ● Read the following: <br><br>1. Precision_and_recall <br><br>**Precision** is the fraction of relevant instances among the retrieved instances <br>**Recall** is the fraction of relevant instances that were retrieved <br><br>2. How-to-properly-evaluate-speech-to-text-engines |

**Some general guidelines recommended for the STT evaluation:**

We should know the target use case for the future voice solution. Then to evaluate speech recognition engines, we need to have representative data related to our use case, with key factors like devices used, environment (eg. noisy warehouse) and accents. Once we have our data, we need to define a test methodology based on our use case. *Never rely ONLY on WER to evaluate the speech engines.* Then Experiment and Evaluate All Features Available.

3. Evaluating an automatic speech recognition service
   **Six steps for performing an evaluation:**
   i.   Choose a test sample
   ii.  Create reference transcripts
   iii. Run a test
   iv.  Create ASR hypothesis transcripts
   V.   Calculate WER
   vi.  Make an assessment

4. N-gram

   An **n-gram** is a contiguous sequence of *n* items from a given sample of text or speech. The items can be phonemes, syllables, letters, words or base pairs according to the application.
   An *n*-gram of size 1 is referred to as a "unigram"; size 2 is a "bigram" (or, less commonly, a "digram"); size 3 is a "trigram".
   English cardinal numbers are sometimes used, e.g., "four-gram", "five-gram", and so on.

5. Text Summarization Evaluation - BLEU vs ROUGE

   **Bleu measures precision**: how much the words (and/or n-grams) in the *hypothesis* appeared in the reference text.
   **Rouge measures recall**: how much the words (and/or n-grams) in the *hypothesis* appeared in the reference text.

| | |
|---|---|
| | 6. [BLEU](#)<br><br>**BLEU** (bilingual evaluation understudy) is an algorithm for evaluating the quality of text which has been machine-translated from one natural language to another.<br><br>7. [The Ultimate Performance Metric in NLP](#)<br><br>**ROUGE-N** measures the number of matching 'n-grams' between our model-generated text and a 'reference.<br><br>**ROUGE-L** measures the longest common subsequence (LCS) between our model output and reference<br><br>● Evaluated the performance of chosen STT package (Watson) over different metrics, namely WER, ROUGE-l, and BLEU and reported the results.<br><br>*Github link:* [https://github.com/ashishguptap99/Speech-to-Text---Watson/blob/main/Unscripted_Ashish__STT_Evaluation.ipynb](https://github.com/ashishguptap99/Speech-to-Text---Watson/blob/main/Unscripted_Ashish__STT_Evaluation.ipynb) |
| 21-27 June 2021 | ● Narrowed down the problem statement and came up with three potential ideas:<br><br>    1. Journalism<br>    2. Meeting minutes<br>    3. Lecture notes<br><br>● Did research on the source codes and datasets for the given use cases and prepared the doc of found resources:<br><br>[https://docs.google.com/document/d/1fsofQJewU3Mdqv2iagi7DSdqBkNh_ryeAUtrcFUqwZw/edit?usp=sharing](https://docs.google.com/document/d/1fsofQJewU3Mdqv2iagi7DSdqBkNh_ryeAUtrcFUqwZw/edit?usp=sharing)<br><br>● Prepared a tentative plan of action for the assigned use case (Meeting Minutes). |

| 28 June - 04 July 2021 | • Spent around 6 hours on implementation of Text Summarization codes, but failed to run most of them. Finally found some codes which could be run.<br><br>https://colab.research.google.com/drive/17jdoJGa_JTHV8R4QNBYLm4B3CJjsm7t-?authuser=1#scrollTo=gPh1of0PG4Ky<br><br>• Also tried to implement the speech_recognition on a large audio file. (1 Hour)<br><br>https://colab.research.google.com/drive/1BjVVvKH7W50c_ogtKkEthZrCoHQqPwMW?authuser=1#scrollTo=Gy7RVyR6vgad<br><br>• Read the 2nd chapter of the book *"Sculpting data for ML"* |
|---|---|
| 05-11 July 2021 | • Tried out the keywords extraction https://colab.research.google.com/drive/11BEQpOAV1uxO1DmJYRJAbJdtz0YTmo0D?usp=sharing |
| 12-18 July 2021 | • Tried STT on our meeting audio recording, but they failed to yield reasonable results on the Indian accent, so we had to go for the extension for speech to text. (2 Hours) |
| 19-25 July 2021 | • Tried to include keywords in meeting minutes, but couldn't be able to get satisfactory results even after exploring various approaches, so decided not to put keywords in the final meeting minutes. https://colab.research.google.com/drive/11BEQpOAV1uxO1DmJYRJAbJdtz0YTmo0D?usp=sharing |