Tanay Dixit | dixittanay@gmail.com
Status:US Citizen
Electrical Engineering | IIT Madras

## EDUCATION

| Program | Institution | CGPA/% | Completion |
|---|---|---|---|
| BTech in Electrical Engineering | Indian Institute of Technology, Madras | **9.45**/10 | 2023 |
| XII | Sri Chaitanya Junior College | 80% | 2019 |
| X | St Mary's School, Pune | 96% | 2017 |

## ACHIEVEMENTS

- Part of the 6 member team to represent IIT Madras at the Prestigious Inter IIT hackathon and was awarded the **gold medal** out of the top 21 competing universities in India. (2021)
- **Ranked 1st** out of more than 200 participants, in the Subex AI Hackathon and was awarded a **cash prize of 45k** (2020)
- Secured All India Rank 341 in JEE Advanced 2019 conducted by IITs (2019)
- Scored **perfect SAT score** of 2400 in SAT 2018 subjects test (Math II , Physics and Chemistry) (2018)
- Secured a Score in Top 10% in National Standard Examination in Physics conducted by IAPT (2019)

## SKILLS

- Programming Languages – **Python**, C, SQL
- Tools - **Git,** LATEX, **MySQL**, VS Code,
- Packages –Scikit-learn, **TensorFlow**, **Pytorch**, Keras, Nltk, NumPy, Pandas, Plotly, Flask
- Operating Systems : **Linux**/Windows

## PROFESSIONAL EXPERIENCE

**Data Analyst Intern** at BlueBarrel Solutions (Houston,Texas) Dec 2020 -Jan 2021
Mentor: Tanmay Chaturvedi

- Received a Letter of Recommendation for my Outstanding work as an Intern
- Analysed more than **10 million rows** of energy consumption data and mined for consumption patterns in order to provide recommendations to power suppliers.
- Developed **Interactive Dashboards** using Flask & plotly and reduced the EDA time by 40%
- Worked on **Hypothesis Tests** to better understand customers HVAC(Heating Ventilation and Cooling) usage patterns

## KEY PROJECTS

https://github.com/tanay2001

1. **State of NLG Evaluation Metrics** Feb 2021 - ongoing
   Guide : Prof. Mitesh Khapra | Robert Bosch Center | IITM
   - Evaluated several NLG metrics to identify pitfalls and cause of poor correlations with human scores.
   - Designed **Adversarial Attacks** to find out which criterias NLG metrics fail to score.
   - Consolidated benchmark scores which will help the community design new and robust metrics
   - Working on getting the **paper published in a top conference ie: *EMNLP findings 2021***

2. **Multilingual Text Classification and Headline Generation** March 2021- April 2021
   Winning solution for the Bridgei2i Hackathon at Inter IIT Tech Meet 9.0 (**github**)
   - Designed a unique and **scalable multilingual** based text classification and summarization pipeline
   - Designed a **phoneme based kim-CNN architecture** which outperformed XLM-R baseline on subdomain text classification.
   - Fine-tuned **mBART** model for improved accuracy on low resource languages like Hinglish
   - **Reduced the inference time by an order of magnitude** and improved model robustness by introducing latent mixing layers
   - Developed an end2end pipeline which is being adopted by Bridgei2i AI team.

3. **NLP with Deep Learning** Nov 2020 – Jan 2021

CS224n course assignments and projects

- Developed a Word2vec model from scratch using **pytorch**
- Trained a NMT model using seq2seq architecture and improved it with **multiplicative attention**
- Trained a NMT subword model using **CharCNN** to improve performance on unseen words
- Worked on Improving model **BERT** performance on **SQuAD 2.0** by adding co-attention layers

4. **Invoice Recognition System** Jan 2021- Feb 2021

Winning solution for the Shaastra Subex AI Hackathon (github)

- Developed an end2end system for automating the invoice verification process.
- Designed an OCR pipeline using Tesseract for extracting text from invoices
- Developed a table structure recognition model using LayoutLM and graphical neural networks.

5. **Driver Monitoring System** June 2020– Nov 2020

Analytics Club | Center for Innovation | IIT Madras

- Built a driver monitoring system to identify rash driving patterns using telemetry data
- Trained an event detection model using **XGBoost** and achieved accuracy of 89%
- Designed a High Frequency Time series Classification model by modifying the **WaveNet** architecture and achieved 94% AUC
- Developed an dashboard using **Streamlit** to analyse our results on **online data**

6. **Analysis of Recommendation Systems** (report) May 2020 - Jun 2020

Guide : Aniket Patil | vRhythms Software Pvt Ltd

- Worked in a team of 4 to analyse various recommendation algorithms to infer which performed best on ranking metrics
- Used real world data set (more than **half a million** purchases) of Amazon products.
- Optimized **collaborative filtering** & **matrix factorization** performance on ranking metrics by 22%
- Designed a Hybrid Autoencoder recommendation algorithm in **python** by leveraging tensorflow
- Analyzed the models performance based on the **Popularity bias** and **cold start** problem using Novelty and Coverage metrics

## COURSE WORK

- **Applied Programming Lab in Python**
- Numerical Methods in C/Python
- Applied Machine Learning
- **Database Management Systems**
- Series and Matrices
- **Probability, Statistics & Stochastic Processes**
- Linear Algebra II
- Principles of Economics

## EXTRA-CURRICULAR ACTIVITIES

1. Sports Activities
    - Conducted Tennis mentorship sessions for UG and PG students
    - Member of Inter -Hostel Tennis Schroeter team
2. Mentoring 6 freshman undergraduates for their academics , extracurricular and mental well being
3. Member of the Analytics Club at IIT Madras