

# Shayan Shafquat

☎ +91-9800113999 ✉ shafquat.shayan@gmail.com in linkedin.com/in/shayan-shafquat 📄 github.com/shayan823

My primary interest lies in **artificial intelligence** and have hands-on experience in **reinforcement learning**, **natural language processing** and **machine learning** projects. I also have a technical experience of **1.5** years working for a ride hailing startup as a **data scientist**. Currently pursuing my long-term goal, of conducting research in the field of **computational neuroscience** I am actively looking out for **research-based roles** as I review the literature and work on independent projects.

## Education

### Integrated MSc. Mathematics and Computing

INDIAN INSTITUTE OF TECHNOLOGY, KHARAGPUR

Jun 2020 | Kharagpur, IN

CGPA: 7.2 / 10.0

## Projects

### Conversion of large-scale cortical models - INCF

GOOGLE SUMMER OF CODE 2022

Mentor: Padraig Gleeson, Ankur Sinha | UCL London

- Verify and improve the original reduced multi compartmental L5 pyramidal cell developed in **NEURON**
- Convert the multi-compartmental cell including channels, morphological and biophysical properties to **NeuroML**
- Validate, visualize, simulate the expected behaviour of the cell model and sharing them on **Open Source Brain**

### Computational Neuroscience - Neuromatch Academy

JANUARY 2022

- Implemented **epsilon-greedy** decision algorithm and a learning rule to solve a **multi-armed bandit** scenario
- Developed a reinforcement learning agent based on **Q-learning** to solve the **cliff walking** problem
- Using **Dyna-Q** solved **Quentin's World** using model-based RL involving activity, learning and planning
- Modeled the **decision-making 2AFC** task in mice incorporating streak of rewards and punishments as features

### iFair - AI and Ethics

FEBRUARY 2020

- Learned a **generalized data representation** preserving **fairness-aware similarity** between individual records
- Developed the combined objective function involving **utility** and **fairness loss** and minimized that using L-BFGS
- Applied the method on two classification tasks of **Census** and **German credit** dataset with **gender**, **age** as protected group resulting in a gain of **3%**, **9%** in **consistency** (fairness) and a drop of **7%** and **1%** in accuracy respectively

### Portfolio Optimization involving System of Linear Interval Equations

JANUARY 2020

Guide: Prof. Geetanjali Panda | Department of Mathematics | IIT Kharagpur

- Designed a problem of **portfolio optimization** involving equations of return, risk and utility with interval parameters
- Investigated and programmed the concepts of **regularity** in interval matrices as a necessary assumption condition
- Used **least squares** to obtain cost function for the problem and minimised that using iterative **gradient descent**

### Hyperpartisan News Detection

COMPETITION: SEMEVAL 2019

- Implemented **Hierarchical Convolutional Attention Networks** for classifying articles on 5 classes of biasness
- Implemented **Empath** model which analyzes articles on **lexical categories** using them for feature engineering

### Comparing India's top 10 Mutual Fund Houses

COMPETITION: INTER HALL DATA ANALYTICS 2018

- Gold winning model comparing performance on the basis of AUM, NAV, Customer retention and other factors
- Used **LSTM** model using past data of NAV and other macroeconomic factors to **forecast** NAV for the next month
- Used **k-means** to cohort the different funds on the three metrics of risk (alpha, beta and standard deviation)

### Rumour detection in tweets

DECEMBER 2018

Guide: Prof. Pawan Goyal | Department of Computer Science | IIT Kharagpur

- Transformed tweets and comments using **one hot encoding** on each character to feed as an input for the **CNN**
- Trained the CNN model with the label as named entities using **entity tagger tool** finally obtaining the pre-final layer
- Classified the tweets by sequentially parsing the pre-final layer of the tweet and comments on it into the **GRU** model

## Work Experience

### ANI Technologies Private Limited (OLA) | DATA SCIENTIST

Bangalore, IN | Sep 2020 – Jan 2022

- **Pay after ride user identification:** Impact: Increased cashless ride by **5%** with no change in default rate i.e. **2.5%**
  - Labelled the trusted base of **0.5M** users based on the last ride reconciliation status or pending days
  - Improved and used feature store API to get user behaviour of **15M+** users prior to their last ride
  - Trained an **ensemble learning** (Balanced Bagging Classifier) on the trusted users with the AUC score of **0.91**
  - Predicted default probabilities for the non-trusted base and automated the pipeline by scheduling a weekly job, whitelisting **8.7M** users in the first iteration by limiting the overall predicted default rate
- **Improvements in peak pricing module:** Impact: Drop in the conversion prediction error from real by **2.2%**
  - Identified **hotspots** in cities to change peak accordingly in the region and its neighbouring geohashes
  - Learned distance based **paramters** to include intrip cabs in supply improving the allocation of cabs
  - Build a fallback model using **heuristic analysis** for peak-pricing to reduce the impact of outages
  - Used **regression** on the spatio-temporal data for cabs to **forecast** peak price for bike category
- **User level pricing:** Impact: **+1.5%** GMV/Bookings and improvements in net completion in the covid days
  - Learned the factor for each user based on their conversion hence improving the demand signal
  - Validated a hypothesis that of abrupt fare in the churn of a regular user using **control-test** group
- **Traffic lights optimisation:** Minimised wait-time for cars in road-network to mitigate congestion
  - Simulated the traffic scenario using **SUMO** framework and in-house cab-pings data (Baseline established)
  - Devised algorithms to optimise congestion locally and globally with fixed and dynamic traffic lights
  - Developed a **Deep Q learning** based **RL agent** on the environment with policies decreasing wait-time by **18.3%**
- **Others:** In-house utility-package (pyutilsds), Baseline model for next basket **recommendation** (Ola-Foods)

### ANI Technologies Private Limited (OLA) | RE- INTERN

Bangalore, IN | May 2019 – Jul 2019

- **Incentive optimisation for partners:**
  - **Feature engineered** each partner's behaviour utilising historical ride-related data of cities Pune and Kochi
  - Implemented **decision tree** on partner's data and incentives predicting their login hours to make **cohorts**
  - **Minimized** burn by **optimizing** the incentive for each of the cohorts with **constraints** on login hours

## Relevant Coursework

- **Offline:**
  - Regression and time series model
  - Programming and Data structure
  - Stochastic Process
- **Online:**
  - Linear Algebra by Prof. Gilbert Strang
  - Natural language processing
- Deep learning Foundations and Applications
- Operation Research
- Computational Statistics
- Computational Neuroscience by University of Washington
- Statistics for Applications by Prof. Philippe Rigollet

## Technical Skills

**Experience with:** Python, C++, R, Git, AWS, SQL, Linux, Bash, Hive, Spark, LaTeX, NEURON, OpenSourceBrain, NeuroML

**Python Packages:** Numpy, Pandas, Scipy, Sklearn, Keras, Scikit-learn, pyNeuroML, LightGBM/XGBoost, NLTK, Tensorflow

**Modelling:** CNN, RNN, GRU/LSTM, Dyna-Q, Gradient Boosting, Transformers, Autoencoders, GANs, Linear Regression, SVM

## Achievements

- Recieved a **pre-placement offer** after the completion of the summer internship at the end of 8th semester
- Among the top 3 teams in the EXL-EQ case study competition, Top 40 teams selected worldwide in SemEval 2019
- Achieved **99.5** percentile score among 0.2M students who appeared in the **JEE** Advance examination
- Completed an **IEEE** certified workshop implementing **Image Processing** techniques and algorithms
- Part of the Gold winning **Football** team in the Inter Hall Sports Championship for the year 2017 and 2019
- Among top **0.1** percentile in 2012 International Assessment of Indian schools conducted by **Macmillan India Ltd.**