Nickil Maveli

Email: nickilmaveli@gmail.com
Mobile: +91 96635 05871

Education

- **The University of Edinburgh** — Scotland, UK
  MSc by Research in Informatics - Institute for Language, Cognition and Computation (ILCC) — *Sept. 2020 – Present*
  Advisor: Prof. Shay Cohen

- **CMR Institute of Technology** — Bangalore, India
  Bachelor of Engineering in Electronics and Communication; CGPA: 7.58/10 — *Aug. 2011 – July. 2015*

Experience

- **Niki.ai** — Bangalore, India
  Senior Software Developer - NLP — *May 2018 - Aug 2020*
  - **Voice Activity Detection**: Developed a real-time pause detection algorithm on speech signals to identify end-of-speech condition which was later integrated on-device. Upto an SNR of 8 dB, the algorithm was able to correctly segment with an accuracy of 91% on different lengths of speech.
  - **Parse Classification**: Designed a probabilistic model capable of generating a binary outcome and intends to classify if an utterance is valid parse or not at every stage in the conversation. Engineered 63 statistical features from parse tree capable of achieving an accuracy of 97% on the holdout test set.
  - **FAQ Bot**: Developed a bot that could provide responses for about 75 FAQ's in regional languages as well as English through a Pooled GRU in conjunction with FastText model. Customer Satisfaction scores increased by 35% leading to better assistance at every dialogue stage.
  - **Code Mixing**: Built a character-level RNN model to classify utterances as 'Hinglish'(blend of Hindi and English) words with minimum lexicon support in combination with a dictionary lookup. This gave an F score of 82% on the holdout test set and helped in identifying such users for better targeted campaigns.
  - **Discovery Platform**: Implemented an algorithm formed on Likelihood and Reliability scores to automatically generate contextual synonyms at n-gram context and augment concept-grammar data so that manual intervention can be avoided. Perplexity metric was minimized upon several iterations.

- **Sociograph Solutions** — Bangalore, India
  Data Scientist - Research and Development — *Sept. 2016 - Feb. 2018*
  - **Timeseries Forecast**: Built a deep learning time series model capable of generating weekly/monthly sales forecasts of products at SKU level with only few thousands of supervised training data. The RMSE metrics were converging after several iterations.
  - **Opportunity Analysis**: Devised a robust machine learning model to predict and optimize various e-commerce metrics like sales conversion rate, customer lifetime value, average order value, shopping cart abandonment rate for retailers that led to 30% decrease in drop off rate.
  - **Image Recommender**: Created an image search engine capable of recommending similar images based on segmentation as a selective search to increase the mean average precision by 45% compared to the benchmark that was available.

- **Hotify.ai** — Bangalore, India
  Junior Data Scientist — *Feb. 2016 - May 2016*
  - **Twitter Mining**: Aggregated tweets from pre-selected news sources to estimate trending hashtags. Naive Bayes algorithm was shown to be most effective when combining both content and contextual data, successfully classifying tweets as containing a trend or not with an accuracy of 75% - 85%.
  - **Data Visualization**: Created Actionable Insights by analyzing how select twitter users through A/B testing react to real-time events by organizing frequencies into temporal buckets to observe the distribution of tweets over time.
  - **Data Consolidation**: Prepared training data of close to 10k samples by carefully examining RSS feeds of news sources and rated each source using factors such as true reach, amplification and network impact.

Selected Projects

- **Sustainable Fishing**: Conceptualized an algorithm that can identify the number and length of fish, and what species those fish are in the videos. This significantly reduced human review time, increased volume of data to manage and protect fisheries.
- **Spectrum Predictor**: Predicted infrared spectra of different chemicals, taking into account the effects of target chemical loading (mass, fill factor, film thickness), target chemical microstructure, and target chemical - substrate interaction. Spectra Correlation ratio of 0.6681 was obtained comparing to ground truth.
- **Music Recommender**: Getting over 30 million tracks, used collaborative filtering based algorithm with matrix factorization and word embedding to generate recommendations. AUC score of 0.7478 was obtained for future unseen unique user-song pair.
- **Speech Recognition**: Used Speech Commands dataset to build an algorithm that understands simple spoken commands having 12 classes. Scored Multiclass Accuracy of 0.91 after implementing 120 log-mel filterbanks for the best model.

Skills

- **Specialties**: Machine Learning, Deep Learning, Reinforcement Learning, Natural Language Processing, Computer Vision, Recommender Systems, Time-series Forecasting
- **Languages**: Python, R, Matlab, Javascript, SQL, HTML/CSS, LaTeX
- **Tools/Frameworks**: Pandas, NumPy, SciPy, Matplotlib, Scikit-Learn, PySpark, NLTK, OpenCV, Keras, PyTorch, TensorFlow, Flask, Django, Docker, Kubernetes, AWS, Jira, Git, Elastic Stack, MongoDB

Achievements

- **Jigsaw Multilingual Toxic Comment Classification**: Secured Top 2% hosted on Kaggle; Won a silver medal.
- **Quality Forecasting in Cement Manufacturing**: Finished 2nd out of 479 competitors hosted on CrowdAnalytix; Won $2,500 as prize money; Winner representing India.
- **Quora Insincere Questions Classification Challenge**: Secured Top 3% hosted on Kaggle; Won a silver medal.
- **NITI Aayog Indic NLP Workshop**: Selected by niki.ai to volunteer for tackling challenges in regional languages concerning NLP.
- **Cold Start Energy Forecasting**: Secured Top 0.85% conducted on DrivenData; Winner representing India.
- **Predicting Poverty**: Placed 6th out of 2310 competitors organized by DrivenData; Winner representing India.
- **Predicting How Points End in Tennis**: Finished 2nd out of 750 competitors hosted on CrowdAnalytix; Won $2,500 as prize money; Featured in Tennis Australia's official press release. [Details]
- **Click Prediction**: Placed 3rd out of 2975 competitors conducted on AnalyticsVidhya; Awarded cash prize of ₹25,000.
- **Funding Successful Projects**: Finished 4th out of 404 competitors organized by HackerEarth.
- **Stack Overflow Ninja**: Among top Stack Overflow users for tags - Python, Pandas, Numpy, Matplotlib with an overall 20000+ reputation score and more than 2 million post views worldwide.
- **Sphere Online Judge**: Ranked under 2000 worldwide out of 0.3 million users.