Nickil **Maveli**

NLP RESEARCH SCIENTIST

10 Crichton Street, Newington, Edinburgh EH8 9AB, United Kingdom

✉ nickilmaveli@gmail.com | ⌂ https://nickil21.github.io | ◯ nickil21 | ≣ nickil-maveli | ▉ nickil21 | ✕ @nickilmaveli | 🎓 Google Scholar | k nickil21

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   *Aug. 2011 - July. 2015*

CGPA: 7.58/10

Work Experience _____

Niki.ai   *Bangalore, India*

SENIOR SOFTWARE DEVELOPER - NLP   *May 2018 - Sept. 2020*

- Developed a real-time pause detection algorithm on speech signals to identify end-of-speech conditions, which was later integrated on-device. Up to an SNR of 8 dB, the algorithm was able to correctly segment with an accuracy of 91% on different lengths of speech.
- 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.
- Developed a bot that could respond to about 75 FAQs in regional languages and English through a Pooled GRU in conjunction with the FastText model. Customer Satisfaction scores increased by 35%, leading to better assistance at every dialogue stage.
- 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.
- Implemented an algorithm formed on Likelihood and Reliability scores to automatically generate contextual synonyms at n-gram context and augment concept-grammar data to avoid manual intervention.

Sociograph Solutions   *Bangalore, India*

DATA SCIENTIST - RESEARCH AND DEVELOPMENT   *Sept. 2016 - Feb. 2018*

- Built a deep learning time series model capable of generating weekly/monthly sales forecasts of products at SKU level with only a few thousands of supervised training data.
- 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 the drop off rate.
- Created an image search engine capable of recommending similar images based on segmentation as a selective search to increase the MAP by 45% compared to the available benchmark.

Hotify.ai   *Bangalore, India*

JUNIOR DATA SCIENTIST   *Feb. 2016 - May 2016*

- Aggregated tweets from pre-selected news sources to estimate trending hashtags.
- 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.
- 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.

Skills _____

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

Selected Publications _____

👥 Conference   * - equal contribution

1. EdinburghNLP at WNUT-2020 Task 2: Leveraging Transformers with Generalized Augmentation for Identifying   *Online (Virtual)*
   Informativeness in COVID-19 Tweets   [WNUT-2020]
   Nickil Maveli
   Proceedings of the 2020 EMNLP Workshop W-NUT: The Sixth Workshop on Noisy User-generated Text

Achievements

2020	**Secured Top 2% Finish (Won a silver medal)** ❯ Jigsaw Multilingual Toxic Comment Classification [🔗]	*Kaggle*
2019	**Secured 2nd place (Won $2,500 as prize money)** ❯ Quality Forecasting in Cement Manufacturing [🔗]	*CrowdAnalytix*
2019	**Secured Top 3% (Won a silver medal)** ❯ Quora Insincere Questions Classification Challenge [🔗]	*Kaggle*
2018	**Secured Top 0.85%** ❯ Cold Start Energy Forecasting [🔗]	*DrivenData*
2018	**Secured 2nd place (Won $2500 as prize money)** ❯ Predicting How Points End in Tennis [🔗]	*CrowdAnalytix*
2017	**Secured 3rd place (Won a cash prize of ₹25000)** ❯ Click Prediction Hackathon [🔗]	*AnalyticsVidhya*
(2016-present)	**Among Top Stack Overflow Users for tags –** `Python`, `Pandas`, `Numpy`, `Matplotlib` **with an overall 30000+ reputation score and more than 5 million post views worldwide** ❯ Stack Overflow Ninja [🔗]	*StackOverflow*

Service

Academic

Reviewer   AmericasNLP 2021, W-NUT 2020 (Shared Task 2), Scipy 2020 (Tutorials, Talk and Poster Presentations)
Volunteer   EACL 2021, ICLR 2021, EMNLP 2020 (Lead Zoom Volunteer)