Nirant Kasliwal

https://github.com/NirantK

nirantk.com

Email Id: hi@nirantk.com

Mobile: +91-7737-887-058

WORK EXPERIENCE

Sundial.so   Bangalore, India

Data Engineering   *October 2021 - Present*

- Built, deployed and documented Internal Systems for onboarding customers across multiple data sources and formats: Redshift, Bigquery, S3 (Parquet/JSONL)

Verloop.io   Bangalore, India

Machine Learning Lead - Focus Area: Conversational AI   *Jan 2019 - April 2021*

- Built and Deployed the state of the art Intent Understanding for Arabic and Indian English dialects
- Reduced need for training data by **20x** for intent understanding by leveraging data augmentation, weakly supervised models and model loss augmentation (e.g. *mixup*). This saved the company **160 man-months** in annotation-time
- Named Entity Recognition and Probabilistic Parser for different locales (e.g. Names in Middle East vs India) and use cases (e.g. customer support vs lead generation)

Soroco India   Bangalore, India

Software Engineer - Focus Area: Deep Learning for Vision in Document Understanding   *Sep 2017 - Dec 2018*

- Implemented cyclic learning rate to speed up **model convergence by 10-50%**; contributed to data augmentation lib which reduced training data needs by **20x** (from millions to hundreds of thousands) for Document Image Segmentation
- Improved Proof of Concept time by **4x** (4 weeks to 1 week) by using AutoML for Text Classification - Built and deployed fastText based toolkit for rapid prototyping and making data sufficiency estimates

Advanced Technologies Lab - Samsung Research Institute   Bangalore, India

Software Engineer (Research) - Focus Area: Machine Learning on Edge Devices   *Jul 2016 - Aug 2017*

- Edge device Machine Learning for Car Driver Safety Scoring on Indian roads, designed to work with low power (12V, 70 HCA) devices and poor Internet (available for less than 20% of the trip duration) connection

PUBLICATIONS

- **Proceedings of the Fourteenth Workshop on Semantic Evaluation**: Bhange, Meghana and Kasliwal, Nirant
- **Book**: Kasliwal, Nirant. Natural Language Processing with Python: Quick Start Guide: Going from a Python Developer to an Effective Natural Language Processing Engineer. Packt Publishing Ltd, 2018.
- Yash Sinha, Prateek Jain, **Nirant Kasliwal**. "Comparative Study of Preprocessing and Classification Methods in Character Recognition of Natural Scene Images." Lecture Notes in Computer Science, Springer, New Delhi, 2015.

OPEN SOURCE CONTRIBUTIONS

Hindi2Vec   *Key Ideas*

State of the Art Language Modeling for Hindi   *Language Modeling, Vectorization, Sub-word Tokenization for LM*

- **First** pre-trained language model for Hindi, trained on Hindi Wikipedia - which achieved **state of the art perplexity**

NLP-Progress   *Key Idea*

Tracking Progress in Natural Language Processing   *Current state-of-the-art for the most common NLP tasks*

- *Maintainer* for a repository which tracks state-of-the-art in NLP Research
- Covers **36 tasks** in English and 6 non-English languages with contributions from **150+ contributors** with over **13.2K** stars

Making the Best of Jupyter   Appreciated by **Nobel Laureate Dr. Paul Romer**

- Internet's most trusted set of Best Practices for Reproducible Science with Interactive Notebooks

- Curated code snippets for exploring image data

Awesome NLP   **Featured in Github.com ML Collection**

- Primary Maintainer for Awesome-NLP, a hand picked list of Natural Language Processing Resources with *75+ contributors*, **9K+ stars**, *1k+ forks*

- Recommended by Stanford Deep Learning Course (CS230) for Literature Review in NLP

Awesome Project Ideas

- Visited by 40K enthusiasts with almost **3K stars** on the Github page

- Curated list of Machine Learning and Deep Learning Projects Ideas, motivating questions and datasets for learners to practice

Natural Language Processing in Python   *Acclaimed NLP 101 for Engineers*

- Unique collection of interactive notebooks which teaches NLP in a code-first manner to promote curiosity

- The collection covers the most common NLP problems, software engineers solve for and share practical solutions, while sharing how to evaluate correctness itself

SELECTED TALKS

- **PyCon 2019 India**: Smart Reply for Chat using Transformer based Language Models to learn Chat-conversations
- **Inaugural Speaker at PyData Bengaluru**:   Invited talk on Sentence Inflection using Dependency Parsing
- **Wingify DevFest**: Invited talk on Transfer Learning for Indic Languages such as Hindi
- **inMobi Tech Talk**: Invited Talk on Transfer Learning using Deep Learning for Natural Language Processing

FELLOWSHIPS & AWARDS

- **Fast AI Part 2 International Fellowship, 2018 & 2019**: International Fellowship for last two years - Advanced Deep Learning Course, FastAI based on *contributions to open source and forum discussions* from over 4K applicants

- **Microsoft AI Challenge - Top 10 Team, 2018**: Top 10 finish globally amongst 2K teams working on Ranking & Search Data from Bing

- **Kaggle Kernel Award for Natural Language Processing- Winner, 2018**: Award for Best Demonstration of Linguistics & Statistical Parser(s) for *Information Extraction, Visualisation and Data Augmentation* with spaCy

- **Future Group Datathon - Runner's Up, 2018**: Won $1K as the Future's Group Datathon focused on Information Extraction, Data Mining and Ranking from 20+ invited from all over the country

- **AI for Education Global Hackweek, 2017**: Awarded the Best Use of IBM Natural Language API from among 4K submissions globally. Built creative tools for School Teachers in India to automatically bring up recent news stories relevant to a topic

- **Schneider Electric APOGEE Innovation Challenge - Winner, 2015**: Elderly Fall Detection system using depth sensing and an off the shelf camera using inbuilt Random Forest in Microsoft Kinect for pose estimation and a SVM classifier finetuned on pose

- **Microsoft code.fun.do - Runner's Up, 2015**: Mixed Reality Canvas for Children wanting to paint on walls

PROGRAMMING SKILLS

Deep Learning & NLP: Proficient in PyTorch, Keras, fastAI, spaCy, NLTK, StanfordNLP

Software: Past Experience in Horizontally Scaled Deployments with Kubernetes, Web Apps for ML Demos with Flask, Async Demos with FastAPI, DVC for Data Versioning, dbt/SQL for data transformation, Dagster/Orchestration for data pipelines