Nirant Kasliwal

nirant.kasliwal@gmail.com | +91 773 788 7058
Github: nirantk | LinkedIn: nirant

SKILLS

PROGRAMMING

Comfortable:
Python Ecosystem
Learner:
PyTorch • fastAI • spaCy • Numpy • Flask • Scikit Learn
Familiar:
TensorFlow • R • C • C++ • zsh/bash shell • LaTeX

EDUCATION & LEARNING

BITS PILANI

M.Sc. (Tech.) Information Systems
2012 - 2016 | Pilani, Rajasthan, India
Relevant Coursework: • Machine Learning • Pattern Recognition • Data Mining

FELLOWSHIP

International Fellow at FastAI Part 2
Spring 2018
International Fellowship to actively participate in the live offering of fastAI's Part 2

PUBLICATION

Published in Machine Intelligence and Signal Processing by Springer
Improved the accuracy of character recognition in natural scene images on the standard Chars74k dataset

- Proposed a classification technique achieving 72% accuracy (state-of-the-art for chosen dataset in 2013) for classifying characters
- Built a basic image processing operations and ensemble machine learning pipeline

EXPERIENCE

SOROCO | Machine Learning Engineer

October 2017 – Present | Bengaluru

- Text Classification toolkit: Built & deployed fastText based toolkit for rapid prototyping across multiple deployments
- OCR: Contributed to a configurable data synthesizer; implemented 1-cycle training policy to speed up network convergence by 10-50%; improved U-net, DenseNet, ResNet variants; ownership of the product shipping suite

SAMSUNG RESEARCH & DEVELOPMENT | Software Engineer (Research)

August 2016 – September 2017 | Bengaluru

- Prototyped algorithms using R which leverage and assess driver behaviour with reference to safety, deployed in C
- Sub-components: Event detection and classification algorithms running on Internet enabled IoT device inside car

OUTSIDE WORK

ONGOING: NLP IN PYTHON

Series of 5-10 Jupyter notebooks for learning

- Written with code examples and programmer-first mindset here
- Exploring several text processing tools such as spaCy, textacy, gensim, and PyTorch
- Key ideas: String matching for spell check, vectorization for similarity, linguistics, ensemble modeling, deep learning for text classification

ONGOING: AWESOME NLP

Maintainer and Lead Contributor

- Featured in Official Github Machine Learning Collection **Awesome-NLP** is a curated list of Natural Language Processing Resources
- 50+ contributors, 6k+ stars, 1k+ forks

HINDI2VEC

State of the Art Language Modeling + new Datasets for Hindi

- Achieved State of the Art Perplexity 46.81 for Hindi using the AWD-LSTM model from from Smerity et al
- Released: Cleaned BBC Hindi data of 4335 documents for text classification and text summarization.
- Released: Pretrained Language Models that you can use in your classification for transfer learning

HACKATHON WIN: QUERY GENERATION

Find recent & relevant news against middle school textbook content

- **Won** the Best use of IBM Watson API (for querying news database) at the Global Opened.ai Hackathon 2017 | Results link
- Query generation using a sequentially mixed content's style elements, and noun phrases