Sriharsh Bhyravajjula

NLP Researcher | Data Scientist | IIIT Hyderabad
http://darthbhyrava.in | s.bhyravajjula@research.iiit.ac.in | +91-9441161582

EDUCATION

MS in Computational Linguistics
Expected May 2019 | IIIT-Hyderabad

B.Tech in Computer Science
Expected May 2019 | IIIT-Hyderabad
Cumulative GPA : 7.14
Dean's List: 2017, 2018

SKILLS

Programming & Scripting
Python • C/C++ • Shell

Data Science
TensorFlow • scikit-learn • PyTorch

Frameworks
Flask • web2py • Django • Jekyll

Version Control
Git • Gerrit

POSITIONS

Speaker, AICC 2018
Invited for a technical talk on NLP at MLR College, Hyderabad.

Teaching Assistant, IIIT-H
Natural Language Processing
(Monsoon '17, Monsoon '18)

Mentor, IASNLP 2017
Guided NLP projects at LTRC's summer school. (Jul '17)

COURSEWORK

Natural Language Processing • Information Retrieval and Extraction • Artificial Intelligence • Statistical Methods in AI • Data Structures and Algorithms • Systems Analysis and Design

EXTRA CURRICULARS

Chief Editor, College magazine
Aug'16 - Aug'17

Co-ordinator, Literary Club
Aug '16 - Aug '17

POETRY

Published
• 9 poems, Sahitya Akademi's Indian Literature (Issues: 275, 286)
• 3 poems, Muse India (Issue: 62)
• 4 poems, Madras Courier (Sep '17)

WORK EXPERIENCE

Klevu Oy | Data Scientist
Creating, implementing and maintaining an ML-based suite of NLP APIs for Klevu's flagship shopping chatbot, which will upgrade this Finnish startup's smart search service for all of their 3000+ e-commerce partners. (May '18 - Jan '18)

Infocrunch | Data Scientist
Worked on ML and NLP approaches to accurately identify individual local sentiment towards political parties in Karnataka from over 100GB of social media data. (Jan '18 - Mar '18)

MaoMao | NLP Engineer
Worked on collecting, indexing and classifying over 100 GB of data scraped using a browser extension, for a p2p topic recommendation engine. (Sep '17 - Jan '18)

Wikimedia Foundation | GSoC Intern
Built a Python-based bot to automate tasks in the "Thanks" extension in MediaWiki as part of my Google Summer of Code project. (Apr '16 - Jul '16)

RESEARCH

Publications (In Progress)
• Modelling multivariate relations between agents in narratives.
• A computational analysis of tone in The Romantics' poetry.

OTHER PROJECTS

Online News Article Popularity
Predicted popularity of online news articles using machine learning methods like PCA with logistic regression for classification. Accuracies for Naive Bayes, Random Forest, Support Vector Machines (SVMs), Neural Nets were between 53% and 85%.

Health Forum Message Classifier
Used Neural Network architectures like LSTMs, CNNs and SVMs to classify messages into categories on a health-forum, with accuracies varying from 55% to 58%.

Wikipedia Search Engine
Developed a search engine for a 61 GB Wikipedia dump using merge sort and secondary indexing, and used retrieval techniques like term-frequency (tf-idf) for ranking.

Factoid QA System over Unstructured Data
Built a factoid question answering system over an unstructured dataset of 150k articles from The Guardian with an accuracy of 74%.

others
Sentiment analysis on 185k words from a football forum • A rudimentary AI/NLP based football assistant • An IRCbot written in Go • An OpenGL 3D game • A Django portal for EERC lab • A Linux Shell written in C • An Android image-recognition app. • A distributed Windows App hosted on Azure.