

Somesh Singh

website | singhksomesh@gmail.com | 9140827410 | LinkedIn | GitHub | Twitter

EXPERIENCE

MIDAS LABS, IIIT DELHI | RESEARCH INTERN

May 2020 – Current | Delhi, India

- Research: Gradient estimation based white-box **adversarial attacks** and black box attacks on Automatic Essay Scoring (AES) systems with customized policies using Tensorflow and Pytorch.
- Analysing the effect of **morphology** on Transformers and how their embedding space is partitioned in correlation with morphology and vocabulary by investigating its knowledge of non-words vs real words. In languages including Basque, Dutch, English, French, Serbian, and Spanish.
- **Supervised by** Prof. Rajiv Ratn Shah (Dept. of CSE & HCD, IIIT Delhi).

MYSMARTPRICE WEB TECHNOLOGY | DATA ANALYST AND SOFTWARE DEVELOPMENT INTERN

May 2020 – July 2020 | Hyderabad, India

- Built Knowledge-based active and passive **recommendation systems** using Bayesian and Non-Bayesian methods using Java, Python, Django and RESTful APIs.
- Built and deployed overall **Social Media Crawling** tools utilizing Docker container using Java, Python and Selenium integrated with cloud and parallel computing.
- **Supervised by** Sitakanta Ray, Co-Founder at Zariance and MySmartPrice.

PROJECTS

DECEPTICONLP | OPEN SOURCE PROJECT, LEAD DEVELOPER

March 2020 - Present

- DecepticoNLP is a Python based PyTorch Wrapper for robustness monitoring and adversarial debugging of NLP models powered by Travis CI.
- DecepticoNLP has implementations of various character & word level adversarial perturbations for text.
<https://github.com/SforAiDL/decepticonlp>

FLAIREDIT | MACHINE LEARNING WEBAPP

April 2020

- Built a complete text classification web-app using Python and PyTorch and deployed it on Heroku which scrapes data from Reddit, pre processes and classifies its flair tag.
URL:<https://github.com/someshsingh22/FlaiReddit-MIDAS>

TEAMELEVEN | COLLEGE PROJECT: BITS F464

- NRR (Net Run Rate) Gaussian Noise Regularised Auto-Encoders for Player Embeddings in Cricket using PyTorch.
- We introduce a regularised autoencoder-based approach to generate on the fly player/team embeddings, using player's form to predict the outcomes of a match and also predict the loss's margin in terms of Net Run Rate (NRR) to get a more confident measure.
- **URL:**<https://github.com/someshsingh22/TeamEleven>

EDUCATION

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

B.E. (HONS.) IN COMPUTER SCIENCE

Expected July 2022 | Goa, India

Cum. GPA: 7.59 / 10

PUBLICATIONS

SEMEVAL-2020

BPGC at SemEval-2020 Task 11:

Propaganda Detection in News

Articles with Multi-Granularity

Knowledge Sharing and Linguistic

Features based Ensemble Learning

SKILLS

PROGRAMMING

3+ years:

Python • C/C++ • Java

0+ years:

Matlab • R •

TECHNOLOGY

Git/Github • AWS • Linux

Tensorflow/PyTorch • Django/Flask •

Numpy/Scikit-Learn/Pandas

COURSEWORK

GRADUATE

Meta Learning

UNDERGRADUATE

Machine Learning (*Teaching Assistant*)

Information Retrieval (*Teaching Assistant*)

Computer Programming (*Teaching Assistant*)

Data Structures & Algorithms

Artificial Intelligence

Foundations of Data Science

Database Systems

Operating System

Object Oriented Programming

SOCIETIES

Events Head, PoR, CTE

Core member, SAIDL

Head of Duties, LRG

Core member, AXON