

# Somesh Singh

☎ (+91) 9140827410 | ✉ singhksomesh@gmail.com | 🏠 someshsingh22.github.io | 📷 someshsingh22 | 🌐 someshsingh22 |  
🐦 @someshsingh22 | 🎓 Somesh

## Education

### Birla Institute of Technology and Science, Pilani

Goa, India

B.E. (HONS.) IN COMPUTER SCIENCE, MINOR IN DATA SCIENCE

Aug 2018 - Present

- **CGPA** - Minor: **9.1/10**, Overall: **8.25/10**
- **Relevant Coursework** - Machine Learning, Deep Learning, Artificial Intelligence, Object Oriented Programming, Data Structure & Algorithms, Database Systems, Applied Statistical Methods, Operating Systems, Computer Architecture, Cognitive Neuro-science, Linear Algebra, Probability and Statistics, Computer Programming, Theory of Computation
- **Teaching Assistant** - CS F441 **Meta Learning**, BITS F464 **Machine Learning**, CS F469 **Information Retrieval**, CS F111 **Computer Programming**, CS F429 **Natural Language Processing**
- **Projects** News Propaganda Detection, Abusive Tweet Classification with Online Learning, Evidence Synthesis in medical text

## Work Experience

### Adobe

Noida, India

MACHINE LEARNING ENGINEER 2

Jan 2022 - Current

- Built **Adobe Express Caption Writer**: <https://new.express.adobe.com/schedule?postId=new> An LLM pipeline for generating and refining social media captions.
- Worked on **Foundational Multimodal Models** on **Human behaviour**, **Long Term memorability**

### CommercelQ

Bangalore, India

DATA SCIENTIST

July 2021 - Dec 2022

- Won Q4 Awards in team **enablers** for our work on digital shelves.
- Advertisement optimization in e-commerce platforms

### MIDAS Labs, IIIT Delhi

Delhi, India

RESEARCH INTERN

April 2020 - Current

- Researched on gradient estimation based white-box **adversarial attacks** on Automatic Essay Scoring (AES) systems with customized policies. **ArXiv**: <https://arxiv.org/abs/2109.11728>
- Worked on generating data free adversaries, investigating dataset bias and worked on post-hoc protection mechanisms against adversarial attacks. **ArXiv**: <https://arxiv.org/abs/2109.12406>
- **Supervised by** Prof. Rajiv Ratn Shah (Dept. of CSE & HCD, IIIT Delhi).

### MySmartPrice Web Technology

Hyderabad, India

DATA ANALYST AND SOFTWARE DEVELOPMENT INTERN

May 2020 - July 2020

- 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.

## Publications

### Measuring And Improving Persuasiveness Of Generative Models

Under Review

OPENREVIEW

2024

Somesh Kumar Singh, Yaman Kumar Singla, Harini S I, Balaji Krishnamurthy

### LLaVA Finds Free Lunch: Teaching Human Behavior Improves Content Understanding Abilities Of LLMs

Under Review

ARXIV

2024

Somesh Kumar Singh, Yaman Kumar Singla, Harini S I, Changyou Chen, Rajiv Ratn Shah, Veeky Baths, Balaji Krishnamurthy

## Large content and behavior models to understand, simulate, and optimize content and behavior

ICLR Spotlight

INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS

Vienna, 2024

Ashmit Khandelwal, Aditya Agrawal, Aanisha Bhattacharyya, Yaman Kumar, **Somesh Singh**, Uttaran Bhattacharya, Ishita Dasgupta, Stefano Petrangeli, Rajiv Ratn Shah, Changyou Chen, Balaji Krishnamurthy

## Long-Term Ad Memorability: Understanding & Generating Memorable Ads

ARXIV PREPRINT

2024

Harini S I\*, **Somesh Singh\***, Yaman K Singla\*, Aanisha Bhattacharyya, Veeky Baths, Changyou Chen, Rajiv Ratn Shah, Balaji Krishnamurthy

## MINIMAL: Mining Models for Data Free Universal Adversarial Triggers

AAAI

ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE

Vancouver, 2022

Swapnil Parekh\*, Yaman Singla Kumar\*, **Somesh Singh\***, Changyou Chen, Balaji Krishnamurthy, Rajiv Ratn Shah

## BPGC at SemEval-2020 Task 11: Propaganda Detection in News Articles with Multi-Granularity Knowledge Sharing and Linguistic Features based Ensemble Learning

COLING

SEM-EVAL-2020, INTERNATIONAL WORKSHOP ON SEMANTIC EVALUATION

Spain, 2020

Rajaswa Patil\*, **Somesh Singh\***, Swati Agarwal

## AES Systems Are Both Overstable And Oversensitive: Explaining Why And Proposing Defenses

ARXIV PREPRINT

2021

Yaman Singla Kumar\*, Swapnil Parekh\*, **Somesh Singh\***, Junyi Jessy Li, Rajiv Ratn Shah, Changyou Chen

## Honors & Awards

2017 **Merit Scholar**, NSEC, Homi Bhabha Centre for Science Education

Mumbai, India

2022 **Q4 Award**, CommerceIQ

Bangalore, India

## Projects

### Code-Soup

Open Source

ADVERSARIAL DEEP LEARNING

September 2021

- code-soup is the python code for the book "Adversarial Deep Learning" written by Dr. Di Jin, Dr. Yifang Yin, Yaman Kumar, and Dr. Rajiv Ratn Shah, and its tutorials. You can use this in conjunction with a course on Adversarial Deep Learning, or for study on your own. We're looking for solid contributors to help.
- It is an easy to understand collection of algorithms and their tutorials in Pytorch
- URL: <https://github.com/Adversarial-Deep-Learning/code-soup>

### DecepticoNLP

Open Source

LEAD DEVELOPER

March 2020

- 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.
- URL: <https://github.com/SforAiDl/decepticonlp>

### FlaiReddit

Independent Project

MACHINE LEARNING WEBAPP

Apr 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 using Transformers.
- The data extractor extracts posts from a wide time period to eliminate the Bias towards some hot topics and compares runtime and implements a minimal bayes classifier.
- URL: <https://github.com/someshsingh22/FlaiReddit-MIDAS>

### Evidencer

Research Project

APPCAIR

July 2021

- Extracting evidence from medical data using Rule Based Engines, NLP, and IR in an online manner

## Symbolic Learning

APPCAIR & TCS RESEARCH

Research Project

December 2020

- Circuit Solver - Used Capsule Networks and SCAE based approaches to learn parts of electric circuit diagrams with an artificially generated dataset and transferred it on real world test data.
- **URL:** <https://github.com/someshsingh22/Circuits>

## College Projects

Self/Team

GITHUB

- **2048-Language:** Translator and Parsor for a gaming language, for the Game 2048 **GitHub**
- **Team-Eleven:** 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, which considers players' previous performance (form) to predict the outcomes of a match with a comparable accuracy (76%). We also predict the loss's margin in terms of Net Run Rate (NRR) instead of win/lose to get a more confident measure.
- **Sherlocked:** Built a text generator using LSTMs, Attention and Transformers with TensorFlow trained on prose like Sherlock Holmes, Harry Potter and poems by John Keats.
- **Snake-AI:** Built A double DQN based self learning snake environment and agent using PyGame and Keras with extensive hyper parameter tuning and reward engineering.

## Tech-Stack

---

<b>Toolkits</b>	Numpy, Scikit-Learn, PyStan, Tensorflow, PyTorch, Linux, MacOS, HTML, CSS, JS, Django, Flask
<b>Programming</b>	DSA, DBMS, OS, OOP
<b>Languages</b>	C/C++, Java, Python, Julia, $\text{\LaTeX}$

## Leadership & Committees

---

2020	<b>Events Head - PoR</b> , Center for Technical Education	Goa, India
2020	<b>Core Member</b> , Society of Artificial Intelligence and Deep Learning	Goa, India
2020	<b>Leader</b> , Language Research Group	Goa, India