

# Guneet Singh Kohli

✉ guneetsk99@gmail.com

in linkedin.com/in/guneetsk99/

🖱 guneetsk.com

🏠 Google Scholar Profile

## EDUCATION

---

**Bachelors of Engineering, Computer Engineering**, Thapar University  
Cumulative GPA: 9.63/10

2018 – 2022 | Patiala, Punjab, India

## PROFESSIONAL EXPERIENCE

---

### GreyOrange

AI Engineer

Jul 2022 – present

- Developed a probabilistic modeling-based Test Harness for Various Bot Task Scheduler Algorithms testing and validation.
- Reduced the testing time from 5 minutes to under 15 seconds for an Average Size Map for a single scenario.
- Introduced improvements in the GreyOrange Launchpad tool for new customers by a reduction in error from 45% to 18%
- Developing objective function and constraints for multi bot systems
- Tech Stacks involved are Java: Optaplanner, Spring, Pytorch, Streamlit, SQL, AirFlow, Grafana, and InfluxDB.

AI Intern (Research and Development)

Jan 2022 – Jun 2022 | Boston, Unites States

- Improved the rack presentations of bots by 25% by testing new heuristic functions to optimize tasks' distribution among warehouse robots.
- Developed simulation software to generate mock problems for testing Task Scheduler made by the AI team of GreyOrange.
- Tech Stack: Java: Spring, Optaplanner, Streamlit, Scipy, Statsmodels, Numpy, Pandas

**Truelytics**, Lead ML Engineer

Oct 2021 – Jan 2022

- Introduction of ML tools into the company's environment by directly working with Eli Gassert, the CTO of the company
- Created benchmarking framework for the organization to perform one-click data analysis and offline model training

## OPEN SOURCE CONTRIBUTION AND COMMUNITY POSITION

---

**Google Summer of Code Fellow**, CERN Switzerland Geant 4 Team

May 2022 – Sep 2022

- Optimized the ML pipeline of fast simulation for the Geant 4 team @ CERN using Kubeflow [Open Source]
- Worked in a team of 3 where I was mentored by 2 CERN Physicists
- Detailed project discussion is presented here [🔗](#).

**CORD.AI**, Core Founding Member, NLP Community Lead

Cord AI is a research community led by a group of students and early career professionals to promote research and professional participation in AI. Our task is to help young students and working professionals enjoy the flavor of research and work on impactful problems. [🔗](#)

**Save the Children (USA)**,

AI for Social Good [Fight against Online Child Grooming]

- Currently leading the technical department of an Online Anti Child Grooming system through which young children can be saved from groomers and predators on platforms like Youtube, Instagram, Twitter, Steam, Discord, etc.
- Working in collaboration with Snap, Save the Children Sweden

**Project Manager and Technical Mentor, Omdena**, AI for Social Good [🔗](#)

**Odia GEN AI**, Lead Scientist

- We are the leading team in India who are developing LLMs for Indic Languages.
- I am currently leading the initiative of creating an evaluation framework for Multilingual LLMs and working on the analysis of their reasoning capabilities.

## RESEARCH EXPERIENCE

---

**Visiting Researcher [Data Study Group, ASDA]**,

Jul 2021 – Jul 2021

Alan Turing Institute and Leeds University

- Using econometrics modeling and Tree-Based ML, we derived statistical relations between weather conditions and their effects on ASDA Stores LTD, United Kingdom sales.
- A detailed publication of the project can be found here

**Junior Researcher**, Charles University, Czech Republic

May 2021 – Jul 2021

Researched Meeting Summarization and Argument pair mining under the guidance of Dr. Tirthankar Ghosal

**Researcher**, AI-NLP-ML Group, IIT Patna

May 2022 – present

Working on Peer Reviews Mining and Scholarly Document Summarisation.

## PUBLICATIONS

---

**HaVQA: A Dataset for Visual Question Answering and Multimodal Research in Hausa Language**,

Findings of the Association for Computational Linguistics, ACL 2023 [🔗](#)

**TransLearn: A Clustering based Knowledge Transfer Strategy for improved Time Series Forecasting,**  
Knowledge-based Systems Elsevier [IF of 8.139] [↗](#)

**Adversarial Perturbations Augmented Language Models for Euphemism Identification,**  
The Third Workshop on Figurative Language Processing, EMNLP 2022 [↗](#)

**Racoons at SemEval-2022 Task 11: Leveraging Concatenated Word embeddings for Named Entity Recognition,**  
The 16th International Workshop on Semantic Evaluation collocated with NAACL 2022 [↗](#)

**ARGUABLY @ AI Debater-NLPCC 2021 Task 3: Argument Pair Extraction from Peer Review and Rebuttals,** NLPCC,2021 [↗](#)

**Team AINLPML @ MuP in SDP 2021: Scientific Document Summarization by End-to-End Extractive and Abstractive Approach,** Proceedings of the Third Workshop on Scholarly Document Processing, COLING 2022 [↗](#)

**ARGUABLY@SMM4H'22: Classification of Health Related Tweets using Ensemble, Zero-Shot and Fine-Tuned Language Model,** SMM4H @ COLING 2022 [↗](#)

**Other Publications available at Google Scholar** [↗](#)

## ORGANIZATIONS

---

**NYU + Dreamspace + Omdena,** Lead ML Engineer

- The project aimed to create an AI tool to detect Tamil Hate speech on platforms like Facebook and Youtube.
- I was the task lead for the data annotation and modelling phase of the project.

**Save the Children(USA) + Omdena,** Lead ML Engineer [↗](#)

Developed an Anti-Grooming AI Solution that analyses conversation flow and performs contextual sentiment analysis to protect young users from online child abusers, capable enough to mimic both child or groomer-type conversation.

**WeedBot(Latvia) + Omdena,** Lead ML Engineer [↗](#)

Deployed the Yolact ++ instance segmentation architecture on Nvidia Xavier Jetson for detection of weeds growing in the field and exterminating them using a laser system

**Kanda Weather(Africa)+ Omdena,** Lead ML Engineer [↗](#)

Built a 12-hour Rainfall Forecasting AI Model to Mitigate Climate Change Variability in West Africa

## PROJECTS

---

**Difesa (A women security app and hardware)** [↗](#)

An android application (Java, Firebase) developed to enhance women security by providing features like anti cut call, voice-based distress commands, continuous location tracking along with a hardware support to also analyze health conditions of the victims and trigger instant responses in the hour of need.

**Crypto Initial Coin Offering (ICO) Analyzer,** Freelance project for Netherlands based Crypto investors [↗](#)

Analyzed recent ongoing ICOs of cryptos to determine success and forecast the expected market cap based on market sentiments, customer social interactions, and attributes that determine the structure of a crypto finance market using advanced predictive analysis and ML algorithms.

## AWARDS

---

**EMNLP D&I Grant,** Empirical Methods in Natural Language Processing

Was awarded the funding of \$750 to present two workshop papers at Abu Dhabi for EMNLP, 2022

**Women Security Android app+hardware support (DIFESA),** Patent Authority of India (Application no 202011015064)

*Filed on 6/04/2020*

**Finalist in HackMIT 2020, MIT(Boston,USA)**

**Merit Scholarship Holder,** Thapar University

Was awarded 15000\$ worth of scholarships based on consistent academic performance throughout my undergraduate.

## SKILLS

---

**Language** (Python, Java, C++, SQL) • **Cloud Services** (AWS, Firebase, GCP)

**Deep Learning Frameworks** (Pytorch, Tensorflow) • **Web Dev Frameworks** (Streamlit, Flask)

**Cloud Frameworks** (Kubeflow , Docker , Kubernetes)

**Libraries** (Pandas, Numpy, Statsmodels, Scipy, Pycaret, Rapids, SciKit Learn, BeautifulSoup)