Karish Grover

+91 99110 54442 | karish19471@iiitd.ac.in | LinkedIn | GitHub | Portfolio | Delhi, India

Education

Indraprastha Institute of Information Technology (IIIT Delhi) — Delhi, India
B.Tech. in Computer Science and Artificial Intelligence; ***GPA: 9.37/10.00*** — *Aug 2019 – Jun 2023 (Expected)*

- **Research interests:** Natural Language Proc., Graph Machine Learning, Deep Learning, Hyperbolic Geometry

Amity International School, Pushp Vihar — Delhi, India
Grade 12^{th} *(Central Board of Secondary Education - PCM);* ***Percentage: 96%*** — *Mar 2018 – Apr 2019*
Grade 10^{th}*;* ***CGPA: 10.00/10.00*** — *Mar 2016 – Apr 2017*

- **Head Boy** of Student Council; Awarded the prestigious Founder's Trophy for exemplary academic performance; Chairperson's special appreciation award, the most prestigious special student award for all-round performance.

Work Experience

LinkedIn — Bengaluru, India
SDE (Machine Learning) Intern (Manager - S.M. Phaneendra Angara) — *May 2022 – Aug 2022*

- Worked as an intern in the **Misinformation and Virality detection team** under LinkedIn AI division. Proposed a novel hyperbolic kernel annular attention mechanism for explainable misinformation detection.
- Hyperbolic manifold for learning representations of the source post, public discourse and the propagation graph.

Coding Ninjas — Delhi, India
DSA Content Creator and Problem Setter Intern (Manager - Navdeep Sandhu) — *Aug 2020 – Nov 2020*

- Data Structures and Algorithms content creator for Python and Java courses at Coding Ninjas. Created detailed and illustrated notes for **entire DSA course**. Wrote editorials for competitive programming problems for over **50** weekly leetcode contest problems. Created programming questions for languages - Python, Java, C++

Publications

- **Karish Grover**, Phaneendra Angara, Md. Shad Akhtar, Tanmoy Chakraborty. **Public Wisdom Matters! Discourse-Aware Hyperbolic Fourier Co-Attention for Social-Text Classification** Advances in Neural Information Processing Systems 34. (**NeurIPS 2022**) [Accepted]
- Agarwal, Rajat, Varun Khurana*, **Karish Grover***, Mukesh Mohania, and Vikram Goyal. **Multi-Relational Graph Transformer for Automatic Short Answer Grading**. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics. (**NAACL 2022**) [Paper] [Code]
- **Grover, Karish**, and Tanishq Goel. **Humor Analysis using Ensembles of Simple Transformers**. Iberian Languages Evaluation Forum (IberLEF) 2021. (**SEPLN 2022**) [Paper] [Code]

Research Experience

The Hong Kong University of Science and Technology (HKUST) — Hong Kong
Undergraduate Researcher (Advised by: Shizhe Diao and Dr. Tong Zhang) — *May 2021 – Feb 2022*
Affiliated Lab*: Statistics and Machine Learning Lab (Website)*

- Proposed a **M**ulti-d**O**main pa**R**aphrase g**E**nerati**O**n dataset with **V**ariable l**E**ngth simila**R**ity, **Moreover**.
- Given a sentence, we want the model to generate paraphrases with different level of **similarity** scores between the generated paraphrase and the source, and different **lengths** of generated paraphrases.
- Curated the **largest paraphrase dataset - 100 Million** (*Domains*: history, politics, movies, sports, tech).

Indraprastha Institute of Information Technology (IIIT Delhi) — Delhi, India
Undergraduate Researcher (Advised by: Dr. Tanmoy Chakraborty and Dr. Md. Shad Akhtar) — *July 2021 – Present*
Affiliated Lab*: Laboratory of Computational Sciences (LCS2) (Website)*

- Working on the task of misinformation detection and graph representation learning using Hyperbolic geometry.
- Proposed a fusion of hyperbolic graph representation learning with a novel Fourier co-attention mechanism in an attempt to generalise the social-text classification tasks by incorporating public discourse.

International Institute of Information Technology (IIIT Hyderabad) Hyderabad, India
Undergraduate Researcher (Advised by: Dr. Radhika Mamidi *May 2021 – Jul 2021*
***Affiliated Lab**: Language Technologies Research Center (Website)*

- Humour detection and rating of Spanish tweets using ensembles of simple transformers and pre-trained models.
- Participated and stood first in the IberLEF 2021 shared task, by developing a model for binary, multi-class and multi-label classification of humour in Spanish tweets.

Other Relevant Projects

ExPoseNet: Expression and Orientation Aware Facial Unmasking [GitHub] [Report] *Jan 2022 – May 2022*

- We develop a model that can generate unmasked images while preserving the facial orientation and expressions.
- A novel GAN-based approach, enhanced with Fourier convolutions to generate a high-resolution unmasked image.
- **Keywords**: *Computer Vision, Fourier Convolutions, Generative Adversarial Networks, Dilated Refinement.*

NewSim: Multilingual News Article Similarity [GitHub] [Report] *Jan 2022 – May 2022*

- A deep learning based model for the task of measuring cross-lingual and multi-lingual news article similarity.
- Two parallel pipelines:- *graph-based* (Multilingual abstract meaning representation for knowledge graph-level news matching) and *text-based* (Multihead attention over multilingual BERT for text-level news matching).
- **Keywords**: *Natural Language Proc., Graph ML, Abstract Meaning Representation, Multilingual BERT.*

From Untruth to Offensive Content on Social Media [GitHub] [Slides] [Talk] *Jan 2022 – May 2022*

- An interactive tool that performs fine-grained analysis and visualize diffusion patterns on real-time Twitter data.
- Classify content into categories like authentic, fake, satire, imposter, manipulated, hatred, and misleading content.
- **Keywords**: *Information Retrieval, Visualization Tool, Software Development.*

Fake News Disambiguation: Why is it Fake? [GitHub] [Report] [Blog] *Jul 2021 – Dec 2021*

- We present *Attention-enhanced Multi-channel Recurrent Convolutional Net*, for explainable fake news detection.
- Extensive ablation studies show that our model outperforms the baseline systems on two benchmarking datasets.
- **Keywords**: *Explainable Machine Learning, Natural Language Proc., Attention-based Highlighting.*

Hierarchical Convolutional Attention Network for Fake News Det. [GitHub] [Report] *Jul 2021 – Dec 2021*

- We present, Hierarchical Convolutional-Attention Network composed of attention-enhanced encoders.
- A CNN-based network to capture the sequential correlation, in unison with word- and sentence-level encoders.
- **Keywords**: *Natural Language Processing, Hierarchical Attention, Deep Learning, Convolutional Attention.*

Awards & Achievements

- **Kartikeya Gupta Memorial Scholarship** 2022 - Only one in college to be awarded scholarship of **100,000 INR** for outstanding all-round performance in academics and co-curricular activities. (Link)
- Conferred with the **Dean's Award for Academic Excellence** for the year 2021/22. (Link)
- Accepted for **MITACS GRI** internship 2022 with fellowship of **500,000 INR**, at Univ. of British Columbia.
- **Chanakaya UG Fellowship** 2022 Awardee - Granted a fellowship worth **100,000 INR**. (Link)
- Conferred with the **Dean's Award for Academic Excellence** for the year 2020/21. (Link)
- Reward of **2,500,000 INR** by **Intel** and **NITI Aayog**, to establish a Tinkering Lab, out of **6000 teams**. (Link)
- **First Position** (Global) at **HAHA@IberLEF2021**, SEPLN Conference shared task. (Link) [Paper]
- Secured All India **99.2 percentile** in **JEE Mains**, 2019 among more than **12,00,000** candidates.
- **Best Project** Mention at YRONS (Young Researchers Of Natural Sciences), **Dubai**, over **200 teams**. (Link)
- **Best Project Award** , out of **1000** Teams, at Odyssey Of The Mind(OOTM), Eurofest, **Slovakia**.
- Only one to stand in the **top 5** in the entire batch for **7 years** (Grade 5 – 11); **Batch topper** for 4 years.

Skills

Programming: Python, Java, Prolog, C, C++

Technologies: Pytorch, Tensorflow, Scikit-learn, Pandas, NLTK, Keras, MySQL

Development: HTML5, CSS4, JavaFX, Java Swing, Javascript, Bootstrap, JQuery

Relevant Coursework: Machine learning, Deep learning, Artificial Intelligence, Natural language processing, Computer vision, Information retrieval, Advanced programming, Graph theory, Semantic webs, Database Management.