B Block, 901
Apex Kremlin
Siddharth Vihar, Ghaziabad, India
2016csz0003@iitrpr.ac.in | mt4descentis@gmail.com
Webpage : descentis.github.io
Github : github.com/descentis
+91-9871562875

Amit Kumar Verma

RESEARCH SCIENTIST, GUVI GEEK NETWORKS

PROFESSIONAL POSITIONS

Research Scientist
HCL (GUVI Geek Network) *Sept '22 - Present*

- Teaching Data Science courses to students from various background.
- Creating courses, labs, projects, and SOTA content for Data Science Students.
- Established Learning & Research Lab in collaboration with IIT Ropar to engage students with research activities in Data Science.
- Built Chatbot for the organisation leaders to analyse the internal dataset for productivity.
- Working on building Personalised AI Assistant to help learners with their course.
- Worked on delivering collaborative courses with IIT Delhi and IIT Roorkee.

Mentor
GUVI Geek Network *May '22 - July '22*

- Teaching Data Science courses to students from various background.
- Teaching Algorithmic Thinking, Introduction to Python, MongoDB, Statistics for Data-Science.
- Worked on Developing Curriculum for Various Courses.

Applied Research Scientist
CoreCLM *July '21 - Dec '23*

- Developed machine learning models to extract the key elements from legal language.
- Developed core functionalities of the product.
- Developed deep learning models for prescription resolution.
- Designed the AI-based semantic search engine for legal languages.
- Designed and deployed document-based chatbot for contract review using LLM.
- Company Link

Content Developer: The Joy of Computing (NPTEL)
Principal Instructor : Dr. Sudarshan Iyengar *Jun '18 - Nov '19*

- Developed majority of the 'Cool Ideas' as course chapters.
- Designed and managed programming assignments for three consecutive semesters.
- The course is one of the most popular courses and is taken by more than 3 Lakh students.
- Course Link

Associate Program Chair: OpenSym-21 & OSS-22
General Chair: Gregorio Robles *Jan '21 - Present*

- Joined as the Associate Program Chair for the premier conference on open collaboration.. - Organised the OpenSym Asia outreach program in April-21.
- Also joined as a part of organizing commettee member for OSS.
- OpenSym Webpage

Assistant Professor, BIMT Gurugram
Department of Computer Science *Aug '16 - Nov '16*

- Responsible for teaching Data Structures and Discrete Mathematics.
- Taken Programming labs in C.
- Developed the Curriculum for DSA.

Indian Institute of Technology Ropar, Punjab, India
Doctor of Philosophy, Computer Science and Engineering,   *Jan' 17 - Aug' 22*
Thesis Title: Aids to Analyse the Knowledge Building Portals
CGPA: 9

National Institute of Technology Meghalaya, Meghalaya, India
Masters of Technology, Computer Science and Engineering,   *June' 14 - Aug' 16*
CGPA: 9.51 (overall)

Guru Gobind Singh Indraprastha University, Delhi, India
Bachelor of Technology, Information Technology,   *Jul' 09 - Aug' 13*
CPI: 71.65 (overall)

Natural Language Processing, Deep Learning, Collective Intelligence, Open Source Software, Social Network Analysis

Simran Setia, Anamika Chhabra, **Amit Arjun Verma**, and Akrati Saxena. "Mediating effects of NLP-based parameters on the readability of crowdsourced wikipedia articles." *Applied Intelligence* (2024): 1-22.

Simran Setia, S. R. S. Iyengar, Anamika Chhabra, **Amit Arjun Verma**, Neeru Dubey (2021). "How well do the students understand the course contents? Assessing comprehension through course videos". *Journal of Computers in Education*, 1-35.

Simran Setia, Sudarshan Iyengar, **Amit Arjun Verma**, Neeru Dubey (2021, September). "Is Wikipedia Easy to Understand?: A Study Beyond Conventional Readability Metrics". In *International Conference on Computational Collective Intelligence* (pp. 175-187). Springer, Cham.

Neeru Dubey, **Amit Arjun Verma**, S. R. S. Iyengar, Simran Setia (2021, September). "Implicit Visual Attention Feedback System for Wikipedia Users". In *17th International Symposium on Open Collaboration* (pp. 1-11).

Amit Arjun Verma, S.R.S. Iyengar, Simran Setia, Neeru Dubey, "An Open Source Toolkit to Parse and Analyze Online Crowdsourced Portals" being reviewed (accepted) in *Journal of Internet Services and Applications* 12.1 (2021): 1-24. **IF-3.9**

Amit Arjun Verma, Neeru Dubey, S.R.S. Iyengar, Simran Setia, "Tracing the Factoids: the Anatomy of Information Re-organization in Wikipedia Articles" to appear in *Proceedings of the Web Conference 2021* (WikiWorkshop): **CORE-A***

Neeru Dubey, **Amit Arjun Verma**, Simran Setia, S.R.S. Iyengar, "WikiGaze: Gaze-based Personalized Summarization of Wikipedia Reading Session" in *Proceedings of the 3rd Workshop on Human Factors in Hypertext (pp. 1-9)* (Part of ACM Hypertext 2020): **CORE-A**

Amit Arjun Verma, S.R.S. Iyengar, Simran Setia, Neeru Dubey "KDAP: An Open Source Toolkit to Accelerate Knowledge Building Research" in *Proceedings of the 16th International Symposium on Open Collaboration (pp. 1-11)* (ACM): **CORE-B**

Simran Setia, S.R.S. Iyengar, **Amit Arjun Verma**, Neeru Dubey "QWiki: Need for QnA & Wiki to Co-exist" in *Proceedings of the 16th International Symposium on Open Collaboration (pp. 1-12)* (ACM): **CORE-B**

Amit Arjun Verma, S.R.S. Iyengar, Nitin Gandhi "WiTPy: A Toolkit to Parse and Analyse

Wikipedia Talk Pages" in *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020 (pp. 535-536)* (System Demonstration): **CORE-A***

Amit Kumar Verma, Nikhil Prakash, "A Grid-based Model for Generating Scale-Free Networks" in *11th International Conference on Communication Systems & Networks (COMSNETS) (pp. 732-736). IEEE, 2019.)*

Amit Kumar Verma, Manjish Pal, "Evolving social networks via friend recommendations" in *11th International Conference on Signal-Image Technology Internet-Based Systems (SITIS), (pp. 379-383). IEEE*

AWARDS &
ACHIEVEMENTS

-Awarded the **Best Paper (Teaser)** for paper presentation at **OpenSym '20**
-**Academic Excellence Award (Gold Medal) 2014-16**, NIT Meghalaya
-Secured a score of **493** in GATE 2016
-Awarded **MHRD GATE Scholarship** for qualifying GATE-2014 with 554 score in Computer Science and Information Technology.

RESEARCH
PROJECTS

Github Link   **An Efficient Method to Store and Access Wikipedia's Revision History for Large-Scale Analysis**
Collaborators: Dr. Sudarshan Iyengar, Neeru Dubey, and Simran Setia   Jun-20 - present

 - Devised a method of making the entire corpus of Wikipedia, inclusive of the edit history, accessible using an improved compression algorithm
 - Developed a deterministic algorithm by reducing the minimization problem into 0/1 Knapsack maximization problem
 - Upto 94% was achieved using the proposed method
 - We are building a library which can work on GPU for a large scale Wikipedia data
 - We are also developing a detaset repository to deploy the compressed Wikipedia dataset over all languages
 - *Paper submitted in* ACM Hypertext 24'

Github Link   **Knowledge Data Analysis and Processing Platform**
Collaborators: Dr. Sudarshan Iyengar, Neeru Dubey, and Simran Setia   May-18 - Jan-20

 - Proposed a standard data representation format (Knol-ML) for storing the dataset of online collaborative knowledge building portals
 - Developed Python-based open source library to parse and analyze the dataset of collaborative KB portals represented in Knol-ML format
 - Library was deployed on pip and is available for public use
 - *Paper was presented at* OpenSym-20

Understanding the factoids rearrangement in Wikipedia articles
Collaborators: Dr. Sudarshan Iyengar, Neeru Dubey, and Simran Setia   May-20 - Dec-20

 - Analyzed the impact of factoids re-arrangement in Wikipedia articles
 - Built a Sequential Model using the sentence similarity measure to understand the crowd involvement in article development
 - We observed a positive correlation between the average sentence similarity and article's quality
 - *Paper was presented at* WWW-21

Github Link   **WitPy: An analysis toolkit for Wikipedia talk pages**
Collaborators: Dr. Sudarshan Iyengar and Nitin Gandhi   Jun '19 - present

 - Developed a parser for creating a structured set of comments from Wikipedia talk pages
 - Working on creating an open source library to analyze the structured talk pages dataset
 - A demonstration version of this paper was presented at JCDL-20

A grid based model for generating scale free networks
Collaborators: Nikhil Prakash   *Jan-19 - May-19*

- Proposed a model for generating scale-free network
- Provided a theoretical analysis of the proposed model
- *Paper was presented in* COMSNETS-19

OTHER ACTIVITIES

-Completed courses in the first year of PhD (worth 16 credits) with a CGPA of 9.
-Held the Teaching Assistant post for various courses (Introduction to Computing (C and Python),
Data Structures and Algorithms, Advanced Algorithms)
-Held the Teaching Assistant post for NPTEL course "NLP with Deep Learning"
-Organised various workshops on "Introduction to Python"

SKILLS

Languages: C, Python, Javascript
Research Tools: PyTorch, Scikit-Learn, Numpy, Langchain, Keras and Tensorflow
Version control and documentation: Git, sphinx

EXTRA INTERESTS

Codechef: Max 3 star rating
Hobbies: Videogames, Competitive Programming, Football

REFERENCES

Dr. Sudarshan Iyengar
Associate Professor, IIT Ropar   *sudarshan@iitrpr.ac.in*

Dr. Abhinav Dhall
Associate Professor, Flinders University   *abhinav.dhall@flinders.edu.au*

Dr. Gregorio Robles
Professor, Universidad Rey Juan Carlos   *grex@gsyc.urjc.es*

Dr. Mukesh Saini
Assistant Professor, IIT Ropar   *mukesh@iitrpr.ac.in*