Dhananjay Kumar

IIT Kharagpur, West Bengal, India
Mobile No.: + 91-7407623322
E-mail: djsingh00@gmail.com
Github: https://github.com/Djasingh
LinkedIn: www.linkedin.com/in/singh-dj

Education

2017-Present **Ph.D. in Text and Network Mining** , *Centre of Excellence in Artificial Intelligence*, Indian Institute of Technology, Kharagpur, India, Course Work (CGPA): 9.00
Thesis Submitted

2017 **M.Tech in Multimedia Information Processing**, *Indian Institute of Technology Kharagpur, India*, M.Tech CGPA : 8.46

Technical Skills

Programming Language	Python
Libraries & Frameworks	NLTK, Spacy, Gensim, Pandas, Numpy, Scikit-learn, SciPy, TensorFlow, PyTorch, Scrapy, Matplotlib
Operating Systems	Linux (Preferred), Windows
Code Editor	Visual Studio Code, Jupyter, VIM

Projects

PhD:

1. **Researcher Influence Prediction using Academic Genealogy Graph**, *In this work, we proposed the task of Researcher Influence Prediction (ResIP) to predict researchers' future influence in an academic field by modelling the researcher academic lineage growth. To address the ResIP, a number of end-to-end deep learning architectures have been proposed.* ,
Skills: LSTM Encoder Decoder, Graph Neural Network, Graph Variational Autoencoder, BERT, Node2Vec, RNN, Mathematics Genealogy Project

2. **Analysis of the Academic Genealogy Graph in an Indian ETD Repository**, *In this work, we studied the academic genealogy network (AGN) in Shodhganga which is the Indian Electronic Theses and Dissertations (ETD) database. We have disambiguated the names of the researchers in Shodhganga and constructed the Shodhganga-AGN, which we have analyzed with genealogical metrics proposed in the literature to unravel several patterns in higher education.* ,
Skills: Text Pre-Processing, Text Similarity, Name Disambiguation, Academic Network Construction, Network Metrics Analysis, Subgraph Extraction, Neo4j Database, Shodhganga Thesis Repository

3. **Tracing the Evolution of Research Topics in an Academic Genealogy Graph**, *In this work, the objective of the study was to investigate the impact of topic drift on the research outcome of a research group.* ,
Skills: Text Pre-Processing, Network Analysis, BERT, Graph Neural Network, Topic Modelling (LDA), Mathematics Genealogy Project

4. **Computational Fact Validation from Knowledge Graph using Structured and Unstructured Information**, *In this work, our primary focus was to leverage unstructured information along with structured ones. Our approach considers finding evidence from Wikipedia and structured information from Wikidata, which helps in determining the validity of the input facts (Triple statements).*,
Skills: Knowledge Graph, TransE Embedding, Solr Indexing, Feature Extraction, Supervised Learning, Random Forest, Naive Bayes, Logistic Regression, KNN, Wikipedia, Wikidata,

M.Tech

1. **Learning to Classify Problems (Questions) into Cognitive Domains**, *We used supervised learning approach to classify problems from different courses of Engineering domains to Bloom'S Cognitive levels.*,
Skills: Text Pre-Processing, Features Extraction, Features Selection, TF-IDF, Mutual Information, Multi-label Classification, Word2vec, Deep Neural Network, Multi label KNN

Area of Research Interest

1 Natural Language Processing
2 Graph Representation Learning
3 Computer Vision
4 Generative AI/Large Language Models

Publications

1 **Dhananjay Kumar**, Plaban Kumar Bhowmick, Jiaul H Paik, Researcher influence prediction (ResIP) using academic genealogy network,Journal of Informetrics,Volume 17, Issue 2,2023, 101392,ISSN 1751-1577, https://doi.org/10.1016/j.joi.2023.101392.

2 **Dhananjay Kumar**, Plaban Kumar Bhowmick, Sumana Dey, Debarshi Kumar Sanyal, On the banks of Shodhganga: analysis of the academic genealogy graph of an Indian ETD repository. Scientometrics 128, 3879–3914 (2023), https://doi.org/10.1007/s11192-023-04728-z.

3 Saransh Khandelwal and **Dhananjay Kumar**. 2020. Computational Fact Validation from Knowledge Graph using Structured and Unstructured Information. In Proceedings of the 7th ACM IKDD CoDS and 25th COMAD (CoDS COMAD 2020). Association for Computing Machinery, New York, NY, USA, 204–208, https://doi.org/10.1145/3371158.3371187.

Certifications

1 Databases and SQL for Data Science with Python, IBM
2 What is Data Science?, IBM
3 Career Essentials in Generative AI, Microsoft and Linkedin
4 Generative AI with Large Language Models, Coursera
5 Complete Guide to TensorFlow for Deep Learning with Python, Udemy
6 Advanced Web Scraping with Python using Scrapy & Splash, Udemy

Extra Curricular Activities

1 Finalist of Machine Learning Hackathon Theme in Techgig Codegladiators 2023.
2 Finalist of International Data Analysis Olympiad Competition 2022.
3 Gate CS 2015 Qualified with AIR 421.