RITIKA PANDEY

(224) 830-6181 | pandey29@purdue.edu | ritika-pandey20.github.io/ | www.linkedin.com/in/pandey-ritika/ | Indianapolis, IN

PhD Candidate| Transforming data into actionable insights | Machine Learning | NLP

EDUCATION

PhD, Computer Science Purdue School of Science | Indianapolis, IN Fall 2020 – Summer 2023
Research- Rewiring police officer training networks to reduce forecasted use of force.
Courses- Intelligent Systems, Computational modeling of epidemics, Survival analysis, Algorithm design, analysis & implementation

MS, Computer & Information Science Purdue School of Science | Indianapolis, IN Spring 2018 – Summer 2020
Courses- Data Mining, Deep Learning, Object-Oriented Programming, Data visualization, Database Systems, Big data Analytics

BTech, Computer Science & Engineering Uttarakhand Technical University | Dwarahat, India Fall 2013 – Spring 2017

TECHNICAL SKILLS

ML Libraries: nltk, spaCy, gensim, networkx, pytorch,scikit-learn, tensorflow, keras, matplotlib, Stanford CoreNLP, sciPy, plotly

Techniques: survival analysis, feature engineering, vectorization, supervised/unsupervised learning, optimization, simulated annealing, transfer Learning, deep learning

Programming Languages: Python, SQL, R, JSL, SAS, d3.js, C++

Databases: MS SQL Server, Oracle, MySQL

Tools: Jupyter, Spyder, SSMS, Heidi SQL, Tableau, Weka, Erwin data modeler, R Studio, Git, AWS, Gephi, R Shiny, JMP Pro

EXPERIENCE

Research Assistant – Machine Learning Purdue School of Science| Indianapolis, IN Feb 2018 – Present

- Design, develop and improve novel machine learning models aimed at social harm & criminal justice applications.
- **Impact:** Investigated role of topic modeling & suggested key metrics (topic coherence, gini coefficient) for detecting crime hotspots allowing for more targeted police intervention.
- **Mentoring:** Guided & collaborated with Undergraduate Research Interns (REU) to analyze Reddit data on insights into modern drug culture & provide tools with potential applications in combating opioid crises.
- **Tools/Stacks:** Python, Text Mining, Graph Mining, Tableau, statistical analysis, data visualization, LDA, NMF.

Data Science Intern – Research & Development Roche | Indianapolis, IN Summer 2021, Summer 2022

- Ideated and applied machine learning techniques to assess additional component for blood glucose system which can be helpful in therapy management for diabetic patients.
- **Modeling:** Built boosted neural network for multi-class classification & performed feature engineering to derive valuable insights for model optimization.
- **Tools/Stacks:** JMP, Python, Boosted Neural Networks, Feature Engineering, JSL, hyperparameter tuning, DoE, data visualization.

IT Intern (Data Analytics) Navient Inc. | Fishers, IN Summer 2019

- Built server based analytical model facilitating prediction of application and chargeback keeping human in the loop.
- **Resolving inconsistencies:** Mined and analyzed server information from various data sources & synchronized it across all platforms.
- Worked closely with application development team & influenced the development trajectory in migrating from spreadsheets to front-end application.
- **Tools/Stacks:** Python, Heidi SQL, SCCM, NEAR (Navient Enterprise Application Repository).

PROJECTS

Officer Pairing to Reduce Use of Force Research Project Jan 2021 – Present

- Constructed network survival model for time-to-event of use of force incidents involving police trainees.
- Introduced a network rewiring algorithm to maximize expected time to use of force events upon completion of training which increases the expected time by 10%.
- **Tools/Stacks:** Cox regression, Random survival forest, simulated annealing, feature engineering, Network analytics.

Homicide Investigation Analysis Research Project Nov 2019 – Jun 2020

- Built knowledge graph-based framework of homicide case chronologies that may aid investigators in analyzing homicide cases & allow for post hoc analysis of the key features that determine whether a homicide is ultimately solved.
- Identified suspect, witness, detective using NER & evidence type using keyword expansion and analyzed the association between network statistics of knowledge graph & homicide solvability.
- **Tools/Stacks:** Word2vec, spaCy, genism, LSTM, CNN, Stanford OpenIE, Tensorflow, matplotlib, Gephi, Random Forest, GLM.

MBTI Personality Detection Coursework Project Oct 2020 – Dec 2020

- Performed personality prediction using machine learning & deep learning techniques that may aid psychologist & private sector in gaining better insights into different personality types of interest & potential hires to better the organization's culture.
- Performed multiclass classification and compared performance across several choice of machine learning and deep learning algorithms.
- **Tools/Stacks:** Python, sentiment analysis, doc2vec, random forest, Convolutional Neural Network (CNN), transfer learning, keras.

Addiction Analysis Research Project May 2018 – Aug 2018

- Obtained data from Reddit and trained a binary classifier which predicts a user's transitions from casual drug discussion forums to drug recovery forums.
- Proposed a Cox regression model that outputs likelihoods of such transitions and found that utterances of select drugs and certain linguistic features contained in one's posts can help predict these transitions.
- **Tools/Stacks:** Python, R, random forest classifier, cox regression, doc2vec, odd ratios, LIWC.

PUBLICATIONS

- **Ritika Pandey**, Jeremy Carter, James Hill, George Mohler, **"Rewiring police officer training networks to reduce forecasted use of force"**, 2023. Under Review. ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '23). Association for Computing Machinery, New York, NY, USA.
- **Ritika Pandey**, P. Jeffrey Brantingham, Craig D. Uchida and George Mohler, **"Building knowledge graphs of homicide investigation chronologies"**, 2020. International Conference on Data Mining Workshops (ICDMW), Sorrento, Italy, 2020. https://doi.org/10.1109/ICDMW51313.2020.00115
- John Lu, Sumati Sridhar, **Ritika Pandey**, Mohammad Al Hasan, and George Mohler, **"Investigate Transitions into Drug Addiction through Text Mining of Reddit Data"**, 2019. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19). Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3292500.3330737
- **Ritika Pandey**, George Mohler, **"Evaluation of crime topic models: topic coherence vs spatial crime concentration"**, 2018. IEEE International Conference on Intelligence and Security Informatics (ISI), Miami, FL, USA, 2018. https://doi.org/10.1109/ISI.2018.8587384

INVOLVEMENT

- Teaching assistant (Database System (Undergrad) Fall 2022, Intelligent System (AI) (Grad) Fall 2019, Fall 2021, Data Science (Undergrad) Spring 2019, Software Design Principles (Undergrad) Fall 2019, Data Structures (Undergrad) Fall 2018)
- Society of Women Engineers (SWE)
- Women in Computer Science (WiCS executive committee)
- Second Helpings
- Girls Inc.