RITIKA PANDEY

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

PhD Candidate| Leveraging research to transform data into actionable insights | Machine Learning | NLP

EDUCATION

PhD, Computer Science Purdue School of Science, IUPUI Aug 2020 – Aug 2023(Expected)
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 | GPA 3.7 Purdue School of Science, IUPUI Jan 2018 – Aug 2020
Courses- Data Mining, Deep Learning, Object-Oriented Programming, Data visualization, Database Systems, Big data Analytics

BTech, Computer Science & Engineering BTKIT, Dwarahat, India Aug 2013 – May 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

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 – Data Science & Machine Learning IUPUI | Indianapolis, IN Feb 2018 – Present
Design, develop and improve novel machine learning models as individual contributor 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
Ideate and apply innovative analytics & machine learning techniques to assess additional component for blood glucose system which can be helpful in therapy management for diabetic patients.
Modeling: Build boosted neural network for multi-class classification & perform feature engineering to derive valuable insights for model optimization.
Tools/Stacks: JMP, Python, Boosted Neural Networks, Feature Engineering, JSL, hyperparameter tuning, DoE, data visualization.

Data Analytics Intern Navient Inc. | Fishers, IN Summer 2019
Built a server based analytical model facilitating prediction of application & chargeback associated with servers 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).

PUBLICATIONS

"Building Knowledge Graph of Homicide Case Chronologies", IEEE ICDM MLLD 2020.
"Investigate Transitions into Drug Addiction through Text Mining of Reddit Data", ACM SIGKDD 2019.
"Evaluation of Crime Topic Models: topic coherence vs spatial crime concentration", IEEE Intelligence & Security Informatics 2018.
"Redditors in Recovery: Text Mining Reddit to Investigate Transitions into Drug Addiction", IEEE Big Data Conference 2018.

PROJECTS

Officer Pairing to Reduce Use of Force Jan 2021 – Present
Constructed network survival model for time-to-event of use of force incidents involving police trainees and 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 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 & 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 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.
Tools/Stacks: Python, sentiment analysis, doc2vec, random forest, Convolutional Neural Network (CNN), transfer learning, keras.

INVOLVEMENT

Teaching assistant, Society of Women Engineers (SWE), Women in Computer Science (WiCS executive committee), Second Helpings, Girls Inc.