

Pankaj Yadav

(315) 742-4170 • pyadav05@syr.edu • [pankajyadav01.github.io/](https://github.com/pankajyadav01) • [linkedin.com/in/pankajyadav015](https://www.linkedin.com/in/pankajyadav015)

EDUCATION

Syracuse University, School of Information Studies, Syracuse, NY Sep 22 – May 2024
M.S. Applied Data Science (4.0/4.0) (Specialization: NLP) Coursework: NLP | Machine Learning | Big Data Analytics Artificial Intelligence Algorithms | Text Mining with LLM | Business Analytics | Statistics | Advance Database Management | Data Science

NorthCap University, CSE Department, Gurgaon, India Aug 15 – May 2019
B.Tech Computer Science and Engineering (8.72/10) Coursework: Linear Algebra | Calculus II | Data Management and Analytics

SKILLS

Programming Languages: Python, R, SQL (MS SQL, MySQL, SQLServer, PostgreSQL), C#, C/C++

Machine learning: Regression, Random Forests, LGBM, XGBoost, Time Series Forecasting, SVM, Decision Trees, kNN

Deep Learning: CNN, RNN, LSTM, Huggingface Transformers, Transfer Learning, Reinforcement Learning, NLP

Big Data: Hadoop, Hive, Spark, Snowflake, DBT

Technologies: AWS, GIT, Jupyter Notebooks, Linux, Docker, Computer Vision, Tableau, PowerBI, ChatGPT

Libraries: TensorFlow, PyTorch, Pandas, scikit-learn, PySpark, NumPy, NLTK, Spacy, ggplot2, dyplr, caret, matplotlib, seaborn

WORK EXPERIENCE

Senior Engineer –Nagarro, Gurgaon, India Jan 2019 – Aug 2022

- Initiated algorithms using various machine learning models and AWS cloud services to extract and segment data from any raw media or document file.
- Led statistical modeling and prediction by using various predictive models, evaluated the accuracy, sensitivity, and specificity for asset categorization and provided the best possible model.
- Implemented algorithms to assess data using Supervised and Unsupervised Learning and other Machine Learning techniques for predicting customer behavior in the Gaming Industry.
- Demonstrated knowledge and application of various cloud services for task achievement and feature deployments

Internet of Things (IoT) Intern, –Ansyst Consulting, Gurgaon, India May 2023 – Jun 2023

- Designed multiple android applications for integrating multiple IoT devices for daily and commercial use.
- Crafted reusable code modules for android apps that can be used for Firebase support to applications.

PROJECTS

Skillspotter: Named Entity Recognition on Job Descriptions – Python, PyTorch, NLU, NLI Sep 2023 – Dec 2023

- Created a dataset of 100K+ rows by web scrapping job portals. Cleaned and tokenized job descriptions for **BERT** model.
- Built a taxonomy of 8000+ soft and tech skills. **IOB tagged** skills using pattern matching and regular expressions.
- Trained a distilbert-base-cased model from **HuggingFace** to identify skills from job descriptions achieving **98% accuracy**.
- Cumulated 34 sets of required skills for different tech roles and built a recommender system based on similarity score.

Yelp Recommendation System – Python, PySpark, Collaborative Filtering Mar 2023 – May 2023

- Spearheaded a team of 4 to develop a recommender system using Yelp customer reviews dataset and Spark.
- Cleaned and transformed 1M+ rows of data followed by feature engineering to implement **K-means and ALS algorithm**.
- Developed scalable framework to recommend 2 similar restaurants for each restaurant and 2 similar users to each user.
- Increased number of relevant recommendations by **60%** through integration of a hybrid **K-means and ALS model**.

Machine Learning based Curation Engine – Python, AWS, Docker Mar 2020 – Mar 2022

- Built a Machine Learning based curation engine that extracts and indexes data from raw media and document file.
- Developed algorithms using Natural Language Processing, Artificial Intelligence, and other Deep Learning Models for predictive data segmentation and extraction.
- Applied multiple Deep Learning techniques to execute Object Detection, Speech Modeling, Video Noise Reduction, Layout Detection, Question Answer Modeling using Natural Language Processing etc.

LEADERSHIP EXPERIENCE

IIT2024 Global Conference, Washington DC – Lead Volunteer Team Jan 12-14, 2024

- Helped organize a team of 50+ volunteers to coordinate 1500+ attendees for a 3-day conference in Washington D.C.
- Managed LinkedIn and Instagram for the conference, growing social media outreach by over 50%.

Senior Engineer, Nagarro – Intern Mentor Mar 2022 – Jul 2022

- Supervised a group of interns and completed multiple projects of high responsibility with clear expectations of showcasing them to prospective future clients