Kirtan Patel

(219)433-8420 | patel.kirtan2004@outlook.com | https://kpatportfolio.vercel.app/

SUMMARY

Data Science student with strong analytical skills and hands-on experience in building scalable data pipelines and predictive models. Proficient in SQL and Python, with a solid foundation in Excel and data visualization techniques for translating complex data into actionable insights. Demonstrated success in collaborative environments, driving data quality improvements and optimizing ETL processes. Eager to apply these skills in analytics and data science to support data management, reporting, and automation initiatives.

EDUCATION

University of Illinois Urbana-Champaign, Siebel School of Computing and Data Science **May 2026**

Bachelor of Science, Data Science

- **Coursework:** Foundation in Data Science, Applied Data Science with Python, Linear Algebra, Probability and Statistics for Computer Science, Data Structures in Python, Intro to Machine Learning, Machine Learning in Physics

EXPERIENCES

Ulta Beauty **May 2025 – Aug 2025**

ML Intern *Bolingbrook IL*

- Designed and implemented **machine learning models** for retail use cases including **personalization**, **forecasting**, and sentiment analysis.
- Optimized data collection, cleaning, and preprocessing pipelines using advanced Excel functions and Python libraries, cutting model training time by **30**%.
- Collaborated with **AI/ML** engineers by writing and refining **SQL** queries to analyze enterprise-scale datasets **(>10M records)**, generating insights that informed **3** new retail AI initiatives.
- Supported brainstorming and research of new AI/ML applications within retail data platforms, identifying opportunities to enhance personalization and **business efficiency**.
- Researched and presented **4** new AI/ML applications for retail platforms, identifying opportunities projected to enhance personalization ROI by **15**%.

Illinois Data Science Club **Aug 2023 - Present**

Project Manager *Champaign IL*

- Built **scalable data pipelines** to process **25GB** of F1 telemetry and weather data, incorporating initial data structuring in Excel alongside **Pandas**, **NumPy**, and **scikit-learn** to ensure high data quality.
- Trained **ML** models using **RandomForest** and **XGBoost** while integrating rigorous data quality checks through grid search optimization for hyperparameter tuning to predict race-winning strategies.
- Developed comprehensive **data visualizations** and interactive **Power BI** dashboards to analyze feature importance, uncover key race-win- ning factors, and support actionable decision-making.
- Designed **predictive pipelines** leveraging ensemble learning models and automated **model evaluation** with cross-validation and custom scoring functions, integrating data quality metrics for improved reliability.
- Secured 2nd place out of 25 teams at IDSC's Data Dive Competition, demonstrating effective team collaboration and commitment to high data quality standards.

PROJECTS

2025 Illinois Statistics Datathon **Mar 2025**

- Built an end-to-end data science solution combining financial forecasting and fraud risk analysis across tokenized and **relational datasets**, incorporating automated data quality validations to guide strategic credit decisions.
- Developed a **chain forecasting model** with linear regression to predict quarterly account spending after performing thorough data quality checks, achieving an RMSE of 1200 despite high variance in user behavior.
- Engineered a RandomForest classifier to detect fraud risk with **86**% accuracy, integrating spending predictions and initial data quality assessments to compute fraud-adjusted credit line recommendations.
- Selected as a finalist out of 180 teams, recognized by Synchrony Financial judges for technical maturity, business alignment, and clear articulation of data quality practices.

Candidate Recommendation Engine **Aug 2025**

- Developed an AI-driven web app that ranks candidates for job descriptions using **fine-tuned SentenceTransformer vector embeddings**, achieving ROC AUC of **0.81** and accuracy of **71**%, significantly improving **semantic matching** accuracy.
- Computed **cosine similarity** between job and resume embeddings to efficiently identify the top 5 most relevant candidates in under 10 seconds per query, enhancing recruitment speed.
- Built a multi-page PDF resume parser capable of correctly grouping resumes for 100% of tested candidates across **50+** files, ensuring flawless **data organization**.
- Integrated Google Gemini **API** to generate automated, **personalized** candidate summaries, highlighting key skills, gaps, and hiring recommendations, reducing manual review time by 40% and boosting productivity.
- Deployed an interactive Streamlit **UI** enabling recruiters to filter, rank, and download candidate summaries efficiently, improving hiring workflow and **user experience**.

SKILLS

- **Languages/Tools**: Python, SQL, R, Java, C++, HTML, Pandas, NumPy, Tableau, AWS, Excel, Azure Cloud, GCP, Power BI
- **Techniques**: Data Preprocessing, Data Wrangling, NLP, Big Data Analytics, Data Mining, Data Pipelines, Data Visualization
- **Other**: Analytical Reasoning, Critical Thinking, Team Collaboration, Problem Solving, Ethical AI, Curiosity