Shruti Pandey

Seattle, WA • (984) 377-9001 • srt.pands@gmail.com • linkedin/shrutipandata • shruti-pandey.com

EXPERIENCE

Data Scientist | ***Nationwide Insurance*** **Jun 2023 - Present**

- Utilized data at Nationwide to drive sales, optimize budgeting, improve business performance, and attain financial objectives through latest innovations in data engineering, business analytics, and machine learning modeling
- Led a data-driven product development to automate 20,000+ annuities decisions per year by deploying a tree-based classifier tailored for imbalanced dataset, utilizing Python, sci-kit-learn, and advanced ensemble techniques with a projected saving of $250,000
- Formulated an advanced Sales Cadence System within a complex data pipeline architecture, leveraging large-scale datasets (2 million+ data points) to revamp wholesaler-advisor interactions, directing lead generation, and cutting acquisition costs by $5 million
- Assessed GenAI (Large Language Model) evaluation frameworks (RAGAS, Llama Index, and HELM) for Nationwide use cases and developed a comprehensive Gen AI governance framework, streamlining LLM integration across business units

Data Scientist, Contract | ***Nationwide Insurance*** **May 2022 - Apr 2023**

- Crafted a linear regression model with dplyr and caret in R, embedded within an ETL data pipeline utilizing SQL for preprocessing, to accelerate life insurance underwriting processes with prescriptive analytics, achieving a 10% reduction in processing time
- Engineered a topic classification system using a Naïve Bayes Multinomial algorithm-based model in Python for categorizing RFP (Request for Proposal) questions for retirement solutions division, attaining 85% accuracy and saving 4.25 hours on each RFP
- Delivered an in-depth analytical review on mortality trends among younger life insurance applicants, using Python and Power BI, synthesizing data from multiple databases and external APIs to inform risk reduction strategies for underwriting and marketing teams

Data Science Consultant | ***Freelance*** **Oct 2019 - Jul 2021**

- Engineered high-impact business solutions for Fortune 500 clients, integrating advanced data pipelines and warehouses; applied Pandas, scikit-learn, and PyTorch for machine learning algorithms, yielding powerful predictive insights
- Crafted a freight pricing strategy by implementing a PyTorch-based deep learning predictive model, integrating API-sourced datasets and sector-specific knowledge, to achieve a 72% accuracy in financial forecasting saving $120,000 in price hedging
- Developed an inventory forecasting model for Tata Motors' auto parts with 83% accuracy by deploying LSTM networks in TensorFlow, collaborating closely with business stakeholders to enhance inventory turnover, achieving an optimal 6:1 turnover ratio
- Designed and deployed dynamic SafeCity dashboard, to visualize live geographic data tracking safety of 1000s of women in real-time, with custom filters using Dash in Python that enriched user interaction, culminating in a 30% uptick in engagement

Associate Data Scientist | ***GSS (a community development organization)*** **Jul 2017 - Sep 2019**

- Facilitated data-driven community development initiatives to improve health and economic outcomes in Indian tribal districts across cross-functional teams by creating and analyzing KPIs using Tableau and Excel to quantify social project impacts
- Designed and implemented a scalable SQL database architecture with automated ETL pipelines and advanced data management practices, overhauling data integrity by 40% and expanding EMBED project's reach to 5,000 additional households
- Led research design, strategic planning, and analytical experimentation, mentoring teams in constructing decisive hypotheses with A/B testing frameworks, and leveraged R, Python, and SQL for comprehensive data analysis to distill actionable insights from field research

Data Science Intern | ***ExxonMobil*** **May 2016 - Jul 2016**

- Delivered market intelligence and forecasted Exxon's 2-wheeler product demand with moving averages, driving strategic decisions

PROJECTS

- Built author classification system with bidirectional LSTM using Pytorch on a dataset created from Project Gutenberg with an accuracy of 68% and performed a qualitative assessment based on GLUE benchmark on generative output by Naive Bayes model
- Coded DeepAR forecasting algorithm (supervised learning algorithm for forecasting one-dimensional time series) from scratch in Pytorch and evaluated its performance on 10-year S&P 500 data, fine-tuned the model to gain a mean RMSE of 0.269
- Collected and engineered EEG data, and devised a hand movement classification model to be included in Brain-Computer Interface with a Support Vector Classifier with an accuracy of 95% in Python

EDUCATION

Master of Science, Data Analytics and Machine Learning | ***Duke University, North Carolina*** Aug 2021 - May 2023

Courses: Linear Algebra, Machine Learning, Deep Learning, NLP, Computer Vision, Advanced Statistics, Usable Security

B.Tech., Mechanical Engineering | ***HBTI, Kanpur, India*** Aug 2013 - May 2017

SKILLS

Technical Proficiency: Programming: Python, R, SQL, NoSQL | Data Science: Machine Learning, Deep Learning, GenAI (LLM) fine tuning and evaluation | Visualization: Power BI, Tableau | Cloud Services: AWS, Azure | Data Warehousing: Redshift, Snowflake

Analytical Expertise: Quantitative Analysis, Applied Mathematics, Statistical Modeling, Predictive Analytics, AI Ethics and Governance

Interpersonal Skills: Business Communication, Leadership, Management, Cross-Functional Teamwork