**FULL NAME**

**(123) 456-7890 | email@mail.com**

**PROFESSIONAL SUMMARY**

Results-driven Data Scientist with 6+ years of experience in deploying machine learning and Generative AI models to drive business impact. Expertise in creating, fine-tuning, and implementing advanced models across production environments, with a proven record in model documentation, monitoring, and adherence to regulatory standards. Proficient in Python, utilizing frameworks like NumPy, Pandas, Scikit-Learn, XGBoost, TensorFlow, and PyTorch. Demonstrated skills in Generative AI, including prompt engineering, Retrieval-Augmented Generation (RAG) models, LLM fine-tuning, and responsible AI practices for bias mitigation and fairness. Strong background in data extraction, transformation, and manipulation, with proficiency in SQL, NoSQL, and big data systems. Known for effective collaboration with cross-functional teams, including Model Risk Management (MRM), IT, and business stakeholders, and adept at selecting modeling techniques to meet both technical and business requirements..

**TECHNICAL EXPERTISE**

* Proven experience in developing and deploying machine learning models in production environments, demonstrating a track record of successful implementation of advanced Generative AI techniques, including fine-tuning and evaluating Large Language Models (LLMs).
* Expertise in machine learning frameworks and libraries such as NumPy, Pandas, Scikit-Learn, XGBoost, Keras, TensorFlow, and PyTorch, with hands-on experience in building robust models for various applications.
* Demonstrated knowledge in prompt engineering and Retrieval Augmented Generation (RAG) models, implementing responsible AI practices to ensure fairness and mitigate bias in model outputs.
* Strong proficiency in statistical analysis and modeling, including methodologies like A/B testing, hypothesis testing, and ANOVA, with experience in time series analysis utilizing ARIMA, SARIMA, and Prophet models.
* Experienced in creating and maintaining detailed model documentation in compliance with industry standards and collaborating with Model Risk Management (MRM) teams for ongoing model monitoring and enhancements.
* Proficient in SQL and skilled in querying structured, semi-structured, and unstructured databases, with extensive experience handling large datasets in Big Data systems and distributed cloud platforms such as AWS, Azure, GCP, and Snowflake.
* Hands-on experience with cloud technologies, including AWS EC2, S3, Azure Data Lake, and Snowflake, implementing scalable and efficient AI/ML solutions across various data formats.
* Practical knowledge of deploying transformer-based architectures like GPT and BERT for generative tasks, focusing on model integration for automated solutions.
* Advanced skills in data visualization and business intelligence, including exploratory data analysis, creating impactful visualizations, and designing BI dashboards using tools like Tableau, Power BI, and Python visualization libraries (Matplotlib, Seaborn).
* Strong communication and collaboration abilities, facilitating effective interactions with technical and non-technical stakeholders, including IT, vendors, and business units to drive model deployment and enhancements.
* Committed to continuous learning and staying abreast of emerging technologies in Generative AI, machine learning, model management to enhance business processes ensuring regulatory compliance.
* Experience working closely with data governance teams and Model Centers of Excellence (CoEs) for ongoing model improvement initiatives.

**TECHNICAL SKILLS & CERTIFICATIONS**

|  |  |
| --- | --- |
| **Languages** | Python, PySpark, JS, HTML, CSS, SQL, HQL, NoSQL, R, |
| **Databases & Cloud Technologies** | MySQL, PostgreSQL, MongoDB, Pinecone, Cassandra, AWS, GCP, Azure. |
| **Big Data & Data Engineering/Streaming** | Apache Spark, Databricks, Snowflake, kafka. |
| **Data Visualization** | Tableau, Power BI, Matplotlib, Seaborn, Plotly. |
| **Data Science/LLM’s Packages** | Spacy, TensorFlow, Pytorch, NLTK, XGBoost, Keras, Pandas, NumPy, SciPy, Scikit- learn, Seaborn, Matplotlib, Hugging Face, OpenAI & LangChain, Ollama. |
| **Machine Learning/DL Techniques** | Supervised Learning, Unsupervised Learning, semi-supervised learning, Deep Learning (DNN, CNN, RNN, LSTM). |
| **Large Language Models/GenAI** | NLP, NLU, LLM’S, fine-tuning, RAG, Llama Series, Genimi pro, Sentiment analysis, Diffusion models, Gan’s, VAE’S, LORA, QLORA. |
| **Deployment Techniques/MLops** | Stream lit, Data button, FastAPI, Amazon SageMaker, AutoML, MLflow, A/B testing, Docker, Kubernetes. |
| **Software Version Control & Documentation** | Git, Jira, Confluence |

**PROFESSIONAL EXPERIENCE:**

**Senior Data Scientist**

**T-Mobile, Bellevue, WA 05/2023 - present**

**Responsibilities:**

* Developed and deployed a chatbot using Python and Generative AI technologies, leveraging Large Language Models (LLMs) on AWS to enhance customer engagement and support efficiency through advanced natural language understanding and generation capabilities.
* Implemented Retrieval Augmented Generation (RAG) models to optimize the chatbot’s response accuracy, employing fine-tuning techniques that ensured high-quality interactions while adhering to responsible AI practices, including bias mitigation and fairness.
* Designed and managed ETL processes on AWS using Python to gather, preprocess, and integrate structured and unstructured data from various sources, including customer feedback platforms, ensuring data consistency and quality for model training and analytics.
* Conducted quantitative analysis using statistical models and machine learning algorithms, such as Random Forest and XGBoost, to assess customer sentiment and enhance the chatbot's ability to understand user intent through effective Generative AI applications.
* Collaborated with the Model Risk Management team and data governance stakeholders to create comprehensive documentation for the chatbot model, ensuring compliance with industry standards and facilitating effective deployment and monitoring of LLM performance.
* Engaged with cross-functional teams, including IT and product management, to align chatbot features with business objectives and user needs, ensuring seamless integration into existing customer service workflows.
* Developed interactive dashboards using Tableau to visualize chatbot performance metrics, user engagement trends, and insights from customer interactions, enabling data-driven decision-making for continuous improvement in the Generative AI space.
* Instituted regular monitoring and evaluation protocols for the chatbot, utilizing user feedback to refine responses and enhance model effectiveness in line with emerging trends in Retrieval Augmented Generation (RAG) technologies.
* Leveraged AWS services such as Lambda for serverless computing and S3 for data storage, ensuring scalability and reliability in handling user queries and managing large datasets efficiently with Python.
* Remained current with advancements in Generative AI and Large Language Models (LLMs) to continually enhance the chatbot's capabilities and ensure compliance with ethical and regulatory standards.

**Environments:** Python, SQL, AWS (Lambda, S3), Tableau, ETL processes, Random Forest, XGBoost.

**Senior Data Scientist** **Verizon, Basking Ridge, NJ 07/2021 - 05/2023** **Responsibilities:**

* Extracted, transformed, and loaded (ETL) data from multiple sources using SQL to develop analytical solutions aimed at predicting customer churn and enhancing customer satisfaction.
* Developed and implemented Machine Learning models using Python and frameworks such as Scikit-Learn, XGBoost, and TensorFlow to predict churn probability, allowing for targeted retention strategies that reduced churn rates by 15%.
* Leveraged Big Data systems like AWS Redshift to analyze customer behavior and churn trends, enabling data-driven decision-making that improved service delivery and customer engagement.
* Deployed machine learning models in production environments using AWS SageMaker, ensuring scalability and efficient processing of large datasets while maintaining high accuracy and reliability in churn predictions.
* Developed and presented interactive Power BI dashboards to visualize churn trends, performance metrics, and predictive insights, enabling stakeholders to monitor key indicators and make informed business decisions.
* Conducted model performance evaluations and continuous monitoring using AWS CloudWatch, ensuring that deployed models met performance benchmarks and provided actionable insights for ongoing improvements.
* Developed comprehensive model documentation aligned with Model Governance standards, collaborating with the Model Risk Management (MRM) team to ensure compliance with industry regulations and best practices.
* Validated model predictions against historical data and industry benchmarks, refining algorithms to enhance predictive accuracy and effectiveness in mitigating churn.
* Integrated analytical outputs into Verizon's customer relationship management systems, providing stakeholders with actionable insights to proactively address customer concerns and improve retention efforts.
* Proficient in SQL, enabling efficient querying and analysis of large datasets to support machine learning initiatives and churn prediction efforts.

**Environments:** Python, SQL, AWS (SageMaker, Glue, Redshift, CloudWatch), TensorFlow, Keras, PyTorch, NumPy, Pandas, Scikit-Learn, XGBoost, Power BI, Jupyter Notebook, Git.

**Data Scientist Cox Communications, Atlanta, Ga 08/2018 - 06/2021**

**Responsibilities:**

* Gathered, interpreted, and manipulated structured data from various sources, leveraging Python and SQL to create comprehensive datasets for analytical solutions. Processed data to enrich customer profiles and enhance model training.
* Developed and validated machine learning models such as Random Forest and Gradient Boosting Machines to improve customer segmentation strategies and enable data-driven marketing initiatives. Utilized ensemble methods like XGBoost for enhanced predictive accuracy.
* Employed feature selection and dimensionality reduction techniques (e.g., PCA, LDA) to optimize model performance, reducing computational costs and processing time on large-scale datasets.
* Developed and implemented clustering algorithms to identify customer segments, allowing for personalized marketing strategies that improved engagement and conversion rates.
* Collaborated with cross-functional teams to design and execute experiments that evaluated the effectiveness of different machine learning techniques in enhancing customer segmentation accuracy, utilizing AWS services for data processing and model training.
* Analyzed large datasets to identify trends and patterns, leveraging SQL for data extraction and manipulation, and utilized insights gained to refine segmentation models and improve marketing campaigns.
* Designed and implemented unsupervised learning techniques to segment customers based on behavioral patterns, enabling targeted marketing and personalized engagement strategies.
* Deployed machine learning solutions using Docker and Kubernetes on AWS, ensuring model scalability and robust integration with existing systems.
* Developed BI dashboards using Power BI to visualize model outputs, enabling business stakeholders to gain insights into customer behavior, trends, and segmentation results.
* Conducted ongoing model monitoring to track model performance in production, collaborating with the Model Risk Management team to ensure continued compliance with regulatory standards.
* Created and maintained comprehensive model documentation as per Model Governance standards, collaborating with the Model CoE to align with business needs.

**Environments:** Python, SQL, Docker, Kubernetes, AWS, Power BI.

**EDUCATION:**

**Master’s in Information Technology – Arizona State University, Tempe, Az.**