Name: Andrew Kate
Contact: (857)-340-3761 | NYC, NY

Summary:
Data-driven Software Engineer with 2+ years of experience in machine learning model development, data preprocessing, and statistical analysis. Skilled in Python, TensorFlow, and scikit-learn to build predictive models, automate workflows, and deliver actionable business insights through data visualization and collaboration with cross-functional teams.

Education:

- Master of Science, Computer Software Engineering, Northeastern University, Boston, MA, 2025 (Expected)
- Bachelor of Management Studies, University of Mumbai, 2019

Skills:

- **Programming:** Python, JavaScript, ASP.NET, HTML, CSS
- **Frameworks & Libraries:** NumPy, Pandas, Scikit-learn, Matplotlib, TensorFlow, ReactJS, NodeJS, ExpressJS, Redux, REST API
- **Concepts & Tools:** Statistics, Probability, Git, Postman, REST APIs, Linux, Bash Scripting
- **Cloud & DevOps:** AWS, GCP, Terraform, Packer, Infrastructure as Code (IaC)
- **Databases:** SQL, MySQL, MongoDB

Experience:
Software Engineer, Yardi Software, Pune, India (Jul 2022 – Jul 2024)

- Designed, trained, and evaluated supervised machine learning models (Logistic Regression, Random Forest, XGBoost) for customer churn prediction, achieving 92% accuracy on validation data.
- Built and maintained data preprocessing pipelines using Pandas, NumPy, and Scikit-learn, improving data consistency across training and production environments.
- Worked with SQL and MongoDB for data retrieval and cleaning, and used Airflow to automate daily model retraining workflows.
- Collaborated with data scientists and product teams to identify ML use cases, define success metrics, and iterate quickly on prototypes.
- Visualized model insights using Matplotlib, Seaborn, and Tableau, enabling stakeholders to understand business impact.

Projects:

- **Machine Learning Model for CO_2 Impact Analysis:** Developed a Python-based ML model to analyze the relationship between rising CO_2 emissions and the frequency of natural disasters. Collected and preprocessed environmental datasets, applied regression and correlation analysis, and visualized trends with Matplotlib.
- **Question Answering NLP Model:** Built a deep learning–based Q&A system using BERT to encode context and question inputs. Implemented preprocessing steps including text cleaning and tokenization, achieving high relevance in automated answers.