PRAMOD PADMAKAR NAGARE

Bellevue, WA **|** +1 857-269-6180 **|** pramodnagare1993@gmail.com

Portfolio: https://pramodnagare.github.io/ **| Linked In:** https://www.linkedin.com/in/pramodnagare/

EDUCATION

MS in Information Systems | Northeastern University, Boston, MA. | Graduation: Dec 2019 | GPA: 3.53/4.0

Major Courses: Advanced-Data Science, Application Engineering & Development, Business Intelligence & Data warehousing, Cloud Computing and Containerization, Cognitive Computing, and Engineering of Big Data Systems

WORK EXPERIENCE

Senior Data Engineer | Institute for Experiential AI, Northeastern University   **Jun 2022 - Present**

- Built data infrastructure and data pipelines for client facing data science projects
- **Led Data Modernization project** for MCGI GTB document extraction infrastructure
- Implemented Movie Recommendation System with **LLM** and **Web 3.0 integration**
- Developed frontend and ML backend for **Cancer and Genomics based internal research projects**
- Conducted technical sessions for the internal team for innovative thinking

Senior Data Engineer | Amgen Inc.   **Nov 2021 – Jun 2022**

- Designed **scalable architecture and Data Lake** solutions for Enterprise-wide Data
- Developed complex data pipelines integrating various data sources across the enterprise
- **Optimized** Databricks pipeline **saving ~ 40% of infrastructure cost and overall runtime 5x improvement**
- Mentored junior data engineers, code review, and drove technical certification drive across the team
- Developed **understanding for HIPPA regulations** for different enterprise research datasets
- **Led onshore and offshore (contractors)** teams with **agile project implementation** and timely deliverables

Software Engineer Lead | Capgemini America, Inc.   **Feb 2020 – Nov 2021**

- Effective collaboration with the Data Science and Business team to understand the use cases with Media industry clients **NBCU and ViacomCBS and, CVS Pharmacy**
- Designed and developed cloud-based scalable architecture on GCP and Azure infrastructure for different clients
- Built multi-terabyte scalable Databricks pipeline in PySpark to drive forecasting data science applications
- Implemented end-to-end CICD and unit testing framework for PySpark development

Machine Learning Engineer Intern| AI Foundry   **Jan 2019 – Aug 2019**

- Designed and **developed deep vision, and, neural network models,** for document image classification
- Implemented a **10x faster and more reliable computer vision solution** for the image processing
- Automated keyword image data collection for classification model with the semi-supervised solution
- Developed a core software algorithm in Python for document key-value extraction and association
- Represented the **development team in Sales meetings** with the clients

Senior Systems Engineer | Infosys Limited, India   **Jul 2014 – May 2017**

- Developed multiple interactive dashboards and reports in MicroStrategy for **Enterprise data analysis**
- Implemented Data Integration ETL workflows from the various data source to Data Warehouse using Talend and SSIS
- Presented **monthly Infrastructure metrics** to clients with **proven strategic solutions** for resource optimization
- Improved report execution with **SQL optimization and process improvement** using automation
- Automated object migration using MicroStrategy tools and VB Script resulted in **saving 1000+ hours** and related costs with **improved customer satisfaction**

CERTIFICATIONS

- **Google Cloud Certified Professional Data Engineer**
- **Databricks Certified Associate Developer for Apache Spark 3.0**
- **Astronomer Certification for Apache Airflow Fundamentals**

TECHNICAL SKILLS

- Cloud Platforms: **Google Cloud Platform, Amazon Web Services, Microsoft Azure, Snowflake**
- Programming Languages: **Python, PySpark, SQL, Java**
- Data Science Packages: **CV, NumPy, Pandas, Scikit-Learn, Matplotlib, NLTK, TPOT, H2O, Keras**
- Business Intelligence Tools: **Tableau, Power BI, MicroStrategy**
- Databases Systems: **Redshift, MySQL, MSSQL, Big Query, MongoDB**
- Big Data Technologies: **Hadoop, Spark, Big Query, HBase, Hive, Pig, HIPPA**
- CICD technologies: **Docker, Kubernetes, Ansible, Jenkins, Circle-CI, GitHub**