Abhijeet Dada Mote

+1 (747) 257-8832 | abhijeetmote@gmail.com | https://www.linkedin.com/in/abhijeet-mote/

Work Authorization: H1B Visa Holder Valid for Another 3 Years, **Current Location**: Dallas, Texas

Professional Summary

Versatile Software Developer with **11+ years of cross-functional experience** spanning **backend Python engineering**, **data engineering**, and **applied machine learning**. Led the design of low-latency, high-throughput systems processing 300K+ requests/sec, built scalable data pipelines for real-time and batch analytics, and deployed ML models into production environments. Primary hands-on expertise in AWS with exposure to GCP and Azure concepts. Skilled in cloud architecture patterns and able to quickly adapt solutions across multi-cloud environments. Adapted to team needs by delivering in roles across software development, data engineering, and data science, especially in AdTech, healthcare, FinTech, and semiconductors. Passionate about clean code, system design, and creating robust, AI-powered, data-driven platforms.

Technical Skill Sets

- **Language:** Python 2.7/ Python3.X, C/C++, Java, shell scripting
- **Testing framework:** Pytest, UnitTest
- **AI/ML Tools & Frameworks**: Scikit-learn, XGBoost, LightGBM, TensorFlow, Hugging Face, Pytorch, YOLO, OpenCV
- **Library**: Pandas, Numpy, SciPy, Dask, Numba, Matplotlib, IPython, Yolo, Boto3, PythonSelenium, Jupyter Notebook, ETL
- **Database**: MySQL, Postgres, MS-SQL, T-SQL, Impala, Oracle, Redshift, MongoDB, Big Query
- **Web technology**: VBScript, javascript, Django, Flask, FastAPI, Tornado, RabbitMQ
- **Operating System**: Ubuntu, Unix, Linux, Centos, AWS Server, SUSE
- **Orchestration Tools**: Docker, Kubernates
- **Visualization Tool**: Matplotlib, PowerBI, Dash-Plotly, Bokeh, Looker, Tableau, Grafana, Streamlit
- **Big Data**: Hadoop, HDFS, Mapreduce, Hive, Pig, Pyspark, Sqoop, Airflow, Glue, Athena, SnowSQL, Snowflake(Beginner)
- **Cloud Platform**: AWS, GCP, Azure, Amazon SageMaker, Google Colab, Zeppelin Notebook, DataProc
- **Data Imputation & Validation:** FancyImpute, DataWig, Pandas Validator, VBA
- **Ticketing Tool**: Jira, Confluence, kanban board
- **Version Control**: GIT, mercurial, SVN
- **CI/CD**: Jenkins, argocd
- **Software Dev practice**: Agile, SCRUM
- **AI Analysis/Debugging:** Jupyter AI, LangChain, Helicone, Gradio
- **AI-Powered Dev Tools:** GitHub Copilot, Cursor, Tabnine, Amazon CodeWhisperer, Kite

Work Experience

MNTN DIGITAL INC **Nov 2022 - Present**

Senior Python Developer / Data Engineer with Data Science & Machine Learning Expertise

Delivering scalable, high-performance backend systems in the Connected TV AdTech space, with a focus on real-time bidding, budget pacing, data pipelines, and anomaly detection.

Python Backend Engineering

- Owned the design and development of a high-performance AdTech budgeting and pacing system from scratch, capable of handling 300–400K requests/sec with 10–100ms response times, leveraging asyncio and FasAPIt/Django REST APIs to maximize concurrency and reduce latency. Also accelerated critical ETL pipelines by 40–50% through parallelization, optimized SQL queries, and improved Spark job configuration.
- Balanced real-time bidding with long-term spend goals using Python, Redis, and PostgreSQL, improving budget utilization, delivery consistency, and system observability. Increased buget efficiency by 25–30% and ensured smoother spend curves across campaigns.
- Collaborated closely with the Bidder team to improve advertisement bidding logic; contributed to the tuning and validation of competitive real-time bidding algorithms.
- Thrived in a fast-paced, feedback-driven environment that values agility and continuous improvement.
- Built a Python decision engine integrated with a high-performance Java/Kotlin bidder, optimizing real-time ad auctions end-to-end.
- Containerized Python services and deployed on Amazon EKS with Helm Charts, ensuring consistent, reproducible, scalable deployments with GitOps-driven CI/CD automation.
- Promoted best practices in Python-based backend development, ensuring code quality, maintainability, and scalability through standardization and clean design principles.
- Designing enterprise applications with OOP principles to build modular, extensible, and efficient solutions.
- Built interactive data apps and dashboards using Streamlit for rapid prototyping and visualization.

Data Engineering/Data Science

- Constructed robust ETL pipelines using Python, PySpark, Databricks, Airflow Scheduler, PostgreSQL, and AWS tools (EMR, Lambda, Athena, CloudWatch, SNS, S3, S3 Glacier, Cost Explorer), enhancing data accuracy and integrity in reporting and analytics through automated validation and anomaly filtering. Incorporated Data Lake and Delta Lake architectures to unify streaming and batch

processing, enable ACID transactions on large datasets, and support scalable analytics.
- Implemented data masking techniques to ensure data privacy and compliance with regulatory standards.
- Used AWS SQS/SNS to forward S3 events to Kafka, enabling real-time ETL and enrichment processing.
- Used Airflow DAG orchestration to streamline multi-stage ETL pipelines across AWS and Databricks.
- Developed PySpark-based solutions for parsing and analyzing large Redis backup files; optimized key-count extraction workflows to enhance downstream processing efficiency.
- Utilized Apache Parquet and ORC file formats extensively within Spark and EMR jobs to achieve efficient, columnar storage, enabling faster reads, predicate pushdown, and reduced I/O in large-scale data processing workflows.
- Applied Agile methodologies throughout the development lifecycle, particularly in creation of the Minimum Viable Product (MVP).Collaborated with cross-functional teams to ensure rapid and iterative development cycle, meeting evolving project requirements
- Architected AWS-based solutions using EKS, S3, Lambda, Athena, and Redshift to support high-throughput data processing, interactive analytics, and real-time anomaly detection. Reduced ETL latency by 50%, cut cloud costs by 25%, and accelerated campaign performance visibility for business teams.
- Implemented cost observability using AWS Cost Explorer, Athena, and custom dashboards**,** helping stakeholders proactively manage pipeline, storage, and compute expenses — leading to a 25% drop in monthly AWS billing.
- Monitored and auto-remediated failures across distributed jobs using CloudWatch and Lambda, improving data pipeline reliability and saving up to 12 hours/month in manual ops intervention.
- Designed interactive analytics workflows using AWS Athena, Glue Catalog, and S3, enabling campaign managers to query terabytes of ad logs in seconds. Reduced manual data retrieval by 90% and improved decision-making on pacing and targeting strategies.
- Optimized cloud costs using FinOps: rightsizing, commitments, automation, Spot, and storage tiering.
- Exploring integration of Snowflake with Python-based ETL pipelines and BI tools for real-time analytics.

AI/ML/GenAI

- Implemented near real-time anomaly detection and ML models for spend and traffic, identifying spikes from misconfigured bids, inaccurate targeting, or bot-driven inventory, preventing budget overspending/underspending.
- Exploring statistical and time-series analysis techniques to detect performance deviations across ads, regions, and publishers.
- Contributed to data visualization and stakeholder-facing analytics efforts, enabling actionable insights on campaign performance and budget pacing, significantly improved how quickly the campaign team reacted to anomalies and underspending.
- Developing Python-based MCP servers and agentic AI workflows, enabling scalable agent interactions with backend systems.
- Prototyping GenAI-powered observability enhancements using OpenAI/Mistral LLMs, LangChain, and function calling to generate natural language summaries of pacing and spend anomalies, leveraging embeddings and cosine similarity to reduce triage time."
- Developing an internal RAG-based knowledge assistant (FAISS, Pinecone, LangChain) that enabled engineers to query domain logic, code structure, and system behavior in natural language, accelerating onboarding and reducing dependency on tribal knowledge.
- Currently exploring and fine-tuning open-source LLMs (e.g., LLaMA, Mistral) from Hugging Face on company-specific data, with ongoing work in applying reinforment learning with Human feedback (RLHF) and multi-agent orchestration.
- Actively exploring how to integrate GenAI into enterprise-scale AdTech environments, focusing on practical, low-risk use cases such as anomaly explanation, knowledge retrieval, and internal tooling support.

Stratus Meridian Inc, USA **Aug 2022 - Nov 2022**

Senior Python Developer

Worked as Senior Python Developer for Client Glorifi Inc, fin tech startup company which provides insurance facility and credit & debit card facility and assistance in mortgage purchase.

- Developed a variety of endpoints/REST APIs using Python FAST API, leveraging Python, PostgreSQL, Azure Functions, and Cosm os DB, which improved data retrieval efficiency for client applications.
- Worked with the frontend team to implement Python logic and used Pydantic for schema validation, enhancing the accuracy and relia bility of data processing in the application.
- Implemented ORM frameworks (SQLAlchemy/Django ORM) to streamline database interactions, reduce boilerplate SQL, and impro ve maintainability of the persistence layer.

UST Global, Malaysia **May 2017 - Jul 2022**

Lead I Python Engineer (Data Engineering Applied ML & Analytics)

Worked as Lead Python Developer, Data Engineer, Automation Engineer, and Data Science Associate for Intel Corporation (USA) and In finity Labs, UST Global Malaysia, delivering automation tools, sensor-based data pipelines, and ML solutions for predictive maintenanc, smart waste management, and anomaly detection. Led a team of 4-6 engineers, mentored in Python, managed project execution, and coll aborated with clients to ensure alignment and successful delivery.

- Designed system architecture and workflows with end stakeholders, enhancing project efficiency and stakeholder satisfaction
- Developed robust, secure, and PEP 8-compliant Python code using Poetry for dependency management and enforced clean code prac tices with Black, flake8, and pylint to ensure efficiency, reusability, and maintainability.
- Performed ETL on diverse data sources and system logs, loaded cleaned data into MS SQL using T-SQL, and created visualizations using Dash, PowerBI, and Matplotlib to identify and interpret patterns in complex datasets.
- Developed Python-based library with Pandas & Dask, enabling data scientists to enhance healthcare data analysis efficiency.
- Developed and deployed a robust web API using Django and FastAPI to manage encrypted clinical patient data, significantly enhancing healthcare data workflow efficiency while ensuring secure and compliant access for clinicians.

- Developed Machine Learning models for classification, regression, and anomaly detection using Pandas, Scikit-learn, and TensorFlow, Supported end-to-end pipelines from feature engineering to predictions, includeing imputation and validation with fancyimpute, datawig
- Improved data processing speed by 30% using PySpark to extract large datasets from HDFS and streamline in-memory computations and ML pipelines.
- Implemented advanced NLP techniques to analyze past issue tracking data, achieving a 40% increase in solution identification efficiency through comprehensive text and data mining.
- Used Yolo's pre-trained object detection library for detecting the objects.
- Worked on the Smart City predictive analytics project and it was nominated for the best project of the year.
- Designed and developed automation workflows and GUI using Python, Pandas, Tkinter, Selenium, Sikuli, and PyAutoGUI.
- Development of the IoT platforms using micro-python & python for sending the sensor data to AWS server, Sage maker for cleaning data using a set of rules, and transforming this data to PostgreSQL, using this set of data for predictions.
- Conducted R&D and developed proof-of-concepts to deliver cost-effective, high-quality solutions, including a PoC using AWS Glue for cleaning sensor data, Built ETL pipelines using Informatica IDMC to integrate AWS and on-prem data.
- Web Scraping of the data using BeautifulSoup, lxml, and Scrapy, performing ELT process on the captured data.
- Development of Perl scripts, migration of Perl to python3
- Worked on Big Data Ecosystem, impala, and MS-SQL database.
- Implementing the performance optimization techniques.
- Development of the Automation Testing framework for field-programmable gate array (FPGA) to avoid human intervention in manual regression testing of FPGA hardware by reducing the time and providing quality solutions using Python, Pandas, and other libraries.
- Worked on IBM Resilient incident response SDK to automate the processes.

Ardentia Technologies Pvt. Ltd, India **May 2015 - May 2017**

Python Developer

Worked on end-to-end development of data processing applications handling ~100 GB of unstructured transactional data from sources such as smart devices and set-top boxes. Delivered two major projects involving transformation of semi-structured data (CSV, XML, JSON, text, XLSX) into standardized XML formats. Actively engaged with clients for requirement gathering and guided junior team members in technical execution.

- Designed and developed scalable ETL pipelines and data processing workflows for semi-structured data; implemented data cleansing APIs using Python, Pandas, lxml, multiprocessing on AWS-EC2
- Visualized trends and anomalies using Matplotlib, PowerBI, and Looker to support business insights
- Implemented and optimized complex stored procedures, conducted performance tuning of Python scripts, and explored open-source alternatives through POCs
- Ensured code quality and maintainability by adhering to PEP-8 standards, using PyLint, standardizing code templates, and documenting with PyDoc to improve readability and reusability
- Wrote and executed unit tests using PyTest and Doctest;
- Migrated legacy Python 2 codebases to Python 3 for better maintainability and long-term support.
- Monitored pipeline health and AWS infrastructure using Nagios and internal health checks; resolved data flow issues and bottlenecks.
- Delivered Python, Git training sessions to new joiners and contributed to knowledge sharing on release management practices.
- Investigated spikes or dips in graphs to identify underlying issues and improve data accuracy.

IIT Bombay, India **Mar 2014 - Feb 2015**

Jr. Software Engineer

Worked on Autograding Project and IITBombayX, OpenEdX in Education domain,

- Developed backend logic and unit tests for the Autograder software; automated workflows using Python and visualized outputs using Plotly Dash. Build internal tools with Tkinter and MySQL.
- Built frontend components using HTML, CSS, and jQuery, improving user interface and experience.
- Automated IITBombayX testing with Selenium and migrated finance workflows from VBA to Python.
- Created interactive learning materials, Python curriculum, assisted in C/C++ workshops under T10KT Program.

Education:

- Master of Computer Applications (MCA-2014) with First Class from Mumbai University
- Bachelor of Information Technology (B. Sc.IT-2011) with First Class from Mumbai University

Extra Activities:

- Reviewed Packt Publication Python Cookbook by Steven F Lott - 2024
- Given training to underrepresented groups in Malaysia 2021 who never done programming.
- Invited as Speaker at PyConX Italy 2019 for giving 3 hours workshop.
- Attended Pycon 2019 in Kuala Lumpur.
- Monthly Python Community Meetup Organizer (Malaysia-Penang 2019-2022)
- Awarded Best Python Trainer in UST Global 2017
- Conducted python workshops for students, corporates & Faculty Development programs.
- Helping Ph.D. candidates using Sunpy Python library for finding flares & intensity of Sun.