Avinash Madhukar Pawar

Bloomington, Indiana, USA
+1 (812)272-0824 | mail.avinashpawar@gmail.com | LinkedIn: avinashmpawar | Github: git-avinashpawar | Portfolio: Avinashpawar.dev

EDUCATION:

Master of Science in Data Science **August 2021 – May 2023**
Indiana University, Bloomington Indiana, USA
Coursework: Statistics, Machine Learning, Cloud Computing, Advanced Database Concepts, High-Performance Computing, Bioengineering.

Bachelor of Technology in Computer Science **June 2016 – March 2020**
Shivaji University, Kolhapur Kolhapur, India
Coursework: Distributed Systems, Operating System, Computer Networking, Database Management Systems, Algorithms, Microprocessors.

SKILLS:

Programming Languages:	Python, Java, JavaScript, C++.
Front End:	HTML, CSS, bootstrap and React library, Redux
Back End:	Flask, Node.JS, .NET
Databases:	MySQL, PostgreSQL, Hadoop, Spark, pgAdmin, JSON and BigQuery.
Version control systems:	Git, GitHub, BitBucket.
Cloud:	Linux, AWS(S3, EC2, Lambda), Cloud native technologies, Docker, Kubernetes.

EXPERIENCE:

Data Analyst (Metadata Content Analyst) **October 2021 – May 2023**
Indiana University, Bloomington Indiana, USA
- Leveraged **SQL** to generate detailed reports for metadata discrepancies, assisting in the identification of unclean or incomplete records.
- Enhanced metadata quality by performing **ETL** processes, involving data cleansing and preprocessing, utilizing **Python** and **Excel**.
- Achieved a remarkable **40%** improvement in metadata quality by implementing systematic data cleansing techniques.
- Expedited metadata processing by **30%** through streamlining data preprocessing workflows and optimizing Excel functions.
- Aided in the development of standardized **metadata** templates, leading to consistency across the library catalog and improving user experience.

Software Engineer **May 2019 - August 2021**
Digital Microsys Technologies, Kolhapur Kolhapur, India
- Designed and maintained **scalable database solutions** for mission-critical applications, ensuring optimal performance and high availability.
- Optimized SQL scripts resulting in a **20%** reduction in query execution time and a **12%** refinement in overall database performance.
- Integrated **RESTful** API web services for precise data retrieval and storage, optimizing external data source interactions.
- Collaborated on developing web applications for a local grocery store and a hotel inventory management system using Django and MySQL. Implemented seamless **e-commerce features** including payment gateway integration, order tracking, and inventory management.

PROJECTS:

COVID-19 Lexicon in Media: An Analytical Perspective | Github
- Directed and led a cross-functional team in the creation of a dynamic **dashboard**, visualizing and analyzing extensive COVID-19 media data.
- Utilized ETL techniques and sophisticated **data pipelines** to process and integrate data from the GDELT dataset, totalling **700 GB**.
- Provided stakeholders an information-rich dashboard to uncover media trends and **patterns**, showcasing expertise in data integration, visualization, and analysis for nuanced pandemic insight. Utilized **GCP**, **BigQuery**, and **Tableau** to translate raw data into valuable insights.

Voice Assistant for Smart Mirror | Github
- Led a team in developing a **voice-controlled assistant** for a Smart Mirror, utilizing Python, JavaScript, SQL, AngularJS, and Bootstrap.
- Developed and deployed Weather, Calendar, Music Player, and News applications on Raspberry Pi with Linux, seamlessly integrating them with the Voice-Controlled Assistant to enhance user interaction and accessibility.
- Integrated an **Android companion app** for remote control and management of the smart mirror's functionalities, enhancing user experience.

Distributed Textbook Search Engine: MapReduce, Cloud Integration, and ETL Pipelines | Github
- Engineered a sophisticated **MapReduce**-based search engine for over **1000** textbooks, integrating ETL pipelines for data acquisition.
- Applied **GCP**, **Node.js** and Google **Cloud Functions** to deploy Mapper and Reducer components, optimizing scalability.
- Built an innovative web interface featuring rapid **sub-second search** results and advanced batch search via file links, streamlining efficiency.
- Showcased versatility in merging cloud deployment, ETL architecture, user-centric interface design, distributed computing, and data engineering.

Sign Language Gesture Classification: Algorithm Exploration and Analysis | Github
- Conducted a comprehensive analysis of the "Sign Language MNIST" project, utilizing various machine learning algorithms for gesture classification. Applied Multinomial **Logistic Regression** with tuned **hyperparameters**, assessing their impact.
- Utilized **SVM**s with various kernels, optimizing hyperparameters, and assessed feature importance through **Random Forest** Classifier analysis.
- Delved into the comparative analysis of machine learning algorithms, optimizing models, and evaluating their performances.

ACHIEVEMENTS & CERTIFICATIONS:

Secretary | Data Science Club at IU October 2021 – May 2023
Google Advanced Data Analytics Professional Certificate | Google | Certificate Link June 2023
Winner**, AWS Game Day challenge |** AWS **|** Indiana Statewide IT Conference April 2023