

# Amol Khade

khadeal@mail.uc.edu ❖ San Jose, California ❖ (513) 652-2242 ❖ [linkedin.com/in/khadeamol](https://www.linkedin.com/in/khadeamol) ❖ [github.com/khadeamol](https://github.com/khadeamol)

## PROFILE SUMMARY

Data professional with 4+ years of experience, with skills and experience in machine learning, computer vision and data engineering. Excellent communication skills and presentation skills, and experience working across business functions. Proficient in modern tech stack for backend and data engineering. **Open to remote work, as well as relocation for on-site.**

## PROJECTS

March 2023 - Present

### Soccer Video Analysis using Computer Vision – Python, OpenCV, YOLO, SuperVision ByteTrack, Roboflow

- Object Detection (Apple Silicon GPU) for player movement using Region of Interest and trackers to create Bounding Boxes
- Used NVIDIA T4 GPU to train custom object detection model using custom Roboflow dataset and YOLO8L
- K-means clustering to extract primary color for each team
- Perspective transformation to account for camera placement, and to calculate camera movement within a clip
- Calculate player speed and distance using object tracking, while accounting for perspective distortion and camera movement
- Calculate ball possession by implementing logic to define ball control and NumPy to calculate distance
- [github.com/khadeamol/OpenCV-YOLO-Soccer-Video/](https://github.com/khadeamol/OpenCV-YOLO-Soccer-Video/)

### Volume Control using Computer Vision – Python, OpenCV, Mediapipe

- Control media volume on their device using just your hand gestures!
- Utilized Mediapipe to detect and visualize hand movements
- Calculate pixel distance between specific hand landmarks, and translate pixel distance into volume levels
- Added text and visual annotations on the output video for volume levels
- [github.com/khadeamol/OpenCV-HandTracking/](https://github.com/khadeamol/OpenCV-HandTracking/)

### Formula 1 Racing Analytics Dashboard – Python, Google Cloud, Streamlit, FastF1, Matplotlib, Seaborn

- Dashboard to analyze telemetry and timing data (millisecond precision) for Formula 1 races ([f1analysisdash.streamlit.app](https://f1analysisdash.streamlit.app))
- Driver Comparison reveals differences in driving styles, lap timing, distance delta, and unravels scope for improvements
- Allows user to understand impact of weather, tire compounds, DRS on lap times, and enable them to optimize strategy

## WORK EXPERIENCE

### xFusion Technologies

February 2022 – March 2023

#### Data Engineer

San Jose, CA

- Developed Python APIs to facilitate data processing and analytics features
- Achieved 80% faster processing on Data APIs by integrating Vaex
- Collaborated with sales and product teams to develop data ingestion, ETL design and development features for the platform
- Custom AirFlow operators for ETL Automation
- **AWS S3 Bucket Explorer:** Python API based service within the platform to upload, access, manage user data stored in AWS S3
- **Data Quality Engine:** Generates statistical metrics and visualizations allowing users a view into quality of their data
- **SQL Query Execution Engine:** Using Dask, DuckDB, this SQL runner allows to query data assets in S3 within the platform

### Avant Financial, LLC

September 2019 – July 2021

#### Data Analyst - Data Engineering Vertical

Chicago, IL

- Saved \$MM by remodeling Credit Card Application database, building checks to ensure optimum data quality and efficiency
- Built Business Intelligence dashboards, reporting mechanisms in **Looker** by identifying & implementing business needs
- Audited ETL processes to identify and implement current and future business use-cases and eliminate bottlenecks
- Improved daily ETL runtime by up to 50% by eliminating bottlenecks and designing a more optimized data model
- Long terms savings of thousands of dollars in projected computation costs through improved ETLs
- Promoted to Data Analyst from Data Management Analyst – expanded responsibilities
- Provided support and bug resolution for ETL processing, demonstrating comprehensive knowledge in programming and data modelling
- Created and owned LookML definitions for Looker dashboards to streamline reporting
- Saved \$\$ cost by automating repetitive reporting requests from business units

### Gloo, LLC

June 2018 - July 2019

#### Data Analyst - Data Science and Strategic Analytics

Boulder, CO

- Automated exploration of new data sources to extend coverage of proprietary Insights platform by 3x
- Performed exploration of novel data source using R scripts (dplyr, tidyr, ggplot2)
- Implemented longitudinal tracking for of data sources through automated R auditing scripts
- Developed **interactive Tableau dashboards** to understand sentiment analysis

## EDUCATION

- **MS Management Information Systems** - University of Cincinnati
- **BE Computer Science** – University of Pune

## TECHNICAL SKILLS

- **Machine Learning and Computer Vision:** OpenCV, PyTorch, Scikit-Learn, TensorFlow, YOLO, SuperVision
- **Python, SQL** (PostgreSQL, NoSQL, SQL Server, MySQL), R (Tidyverse – tidyr, dplyr, ggplot2)
- Extensive experience with Pandas, NumPy, FastAPI, Flask, Boto3, Vaex, PySpark, Streamlit
- **Workflow Orchestration and deployment:** Docker, Kubernetes, Jenkins, Airflow
- **Big data tools:** Spark, Hadoop, Presto, Hive
- **BI and Data Visualization:** Dremio, Alation, Looker, Tableau, Matplotlib, Seaborn
- **Experienced with developing on cloud environments on AWS and GCP**