

# **Soham Khopde**

ML Engineer | Business Data analyst

email: [spkchamp@gmail.com](mailto:spkchamp@gmail.com)

Mob: +91 7977663800

Website: <https://sohamk.github.io/portfolio/>

LinkedIn: <https://www.linkedin.com/in/soham-khopde/>

## **Professional Summary**

Dedicated Machine Learning Engineer / Business Data analyst professional with history of developing multiple machine learning projects, performing thorough data analysis, and proven skills in business analysis.

## **Professional History**

### **Associate Consultant**

**Randstad India Private Limited (KPMG)**

**Mumbai, Maharashtra (On-site)**

01/2025 – 06/2025

#### Responsibilities:

- Supported consulting engagements through data analysis, research, and documentation
- Prepared client presentations, reports, and dashboards using Power BI.
- Assisted in risk assessments, process reviews, and quality control testing.
- Collaborated with cross-functional teams to deliver projects on time
- Ensured compliance with KPMG quality, ethics, and regulatory standards.

Tools Stack: Excel, PowerPoint, SQL, Power BI, Visio, Napkin AI, MS Teams.

### **Business Analyst**

**Nyati Technologies Private Limited**

**Mantralaya,Mumbai,Maharashtra (On-site)**

10/2024 – 01/2025

### **Project: Retrieval Augmented Generation based LLM Chabot**

**Project Description:** A RAG based LLM Chabot which enhances user experience by generating accurate response from large database of government resolution and reduces search time of users and e-governance staff.

#### Responsibilities:

- Conducted Interview sessions with e-governance staff for elicitation of requirements.

## **Education**

- Master of Information Technology(PC)- Data Analytics  
(Testamur - [Link](#))  
GPA: 2 (Credit 60 - 69%)  
2019 – 2020

Swinburne University of Technology  
Hawthorn, Victoria, AU

- Bachelor of Engineering (Computer Engineering)  
CGPI: 6.51  
2015 – 2018

K.C. College of Engineering, Thane

## **Skills**

### Business analysis

- Requirements Gathering & Elicitation (BRD, SRS, FRD, Project Plan & Report)
- Wireframes and Prototyping in Adobe XD
- Process Analysis, Modelling and Improvement
- UML Diagrams & Architecture & process flow diagram in MS Visio & in Figma.
- UAT (Test Cases), Solution Evaluation using KPI and metrics.

### Business Intelligence and Visualization

- SSIS, SSRS, Azure Data Factory
- Power BI, Tableau
- Data Modelling

### Database

- Advanced MS-SQL (Stored Procedures Optimization, Query plan Execution)
- Azure SQL DW

### Data Analysis & Machine Learning

- Multivariate EDA

- Preparing Requirements specification document, process modelling RAG application, preparing LLM prompt templates.
- Preparing Knowledge based for LLM models and implementing solution design for structure of Vector Databases.
- Prototyping Integration of RAG application with retrieval systems, Vector Databases and Generative AI LLM model.

Tools Stack: Ollama server - Llama 3, Jupyter Notebook, Chroma DB, Unstructured IO, MS Visio, LangChain.

#### **Data Analytics Advisor intern**

**Employment Guarantee Scheme (Planning) Department Government of Maharashtra ([https://nrega.nic.in/MGNREGA\\_new/Nrega\\_home.aspx](https://nrega.nic.in/MGNREGA_new/Nrega_home.aspx))  
Mantralaya,Mumbai,Maharashtra**

10/2023 – 09/2024

#### Responsibilities:

- Bid process management, from tender preparation till on-boarding of the vendor, project management and monitoring support for the department.
- Gap analysis of UI/UX and Process improvement of the department's mobile application and website for promoting horticultural initiatives and other demand generation for agricultural products across Maharashtra.
- Stakeholder management and front-end development for department.
- Enhancing dashboard using logical layout design, relevant metrics, effective information presentation and intuitive visualization.
- Implementation of data-driven approach for Management Information Systems (MIS).
- Data Analysis on NREGASoft portal to gain key statistics and provide insights to improve various schemes of department.
- Prepare reports and insights from NREGASoft data and other sources.
- Enable data-driven decision-making by providing relevant, actionable insights to the department.
- UAT and Benchmark the department's portal against official government standards (GIGW 3.0) and provide suggestions for interface and structural changes to enhance overall usability and engagement.

#### Tool Stack:

- Python, My SQL, MS SQL, MS Visio, Adobe XD, Power BI, Jira.

#### **Data analytics Consultant**

**Integrated Learning Systems India Private Limited**

**Remote**

08/2022 – 04/2023

#### Responsibilities:

#### Project 1: Neonatal Sepsis Prediction

- Data Pre- Processing & Cleaning
- Data Imputation, Encoding
- Feature Engineering, PCA
- K-fold Cross Validation, Ensemble Voting Classifier, Neural Networks(MLP Classifier), Hyper-parameter tuning
- Main Libraries - Scikit-Learn, scipy, Pandas, imbalance-learn, lightGBM, Flask.

## **Certifications**

### Business Analysis

- IIBA CBDA



- Exposure to IIBA ECBA standards and techniques

### Business Intelligence

#### Enterprise DNA

- Scenario Analysis-Deep Dive Power BI
- Budgeting and Forecasting
- Advanced Analytics in Power BI
- Financial Reporting with Power BI

### Achievements

- <https://avyantra.com/> - Predictive Scoring Application for detecting sepsis infection in infants below 1 year.

**Project Description:** PreSco is a web application which uses machine learning algorithms to provide for an early risk assessment of Neonatal Sepsis and aids in informed decision making for antibiotics. PreSco is non-invasive and provides a rapid risk assessment of babies.

#### Responsibilities:

- Requirement Gathering (BRD, SRS) and analysis for Predictive Scoring application.
- Conducting Brainstorming sessions for strategic data analysis planning.
- Managing change strategy for data analysis plan and Machine Learning process.
- Developing and improvising Process model for Machine Learning Application
- Designing Wireframes in Figma and creating web application forms using React-MUI and node js.
- Multi-variate Exploratory Data analysis and Model Prototyping in Jupyter Notebook.
- Development of Pipeline process of Data cleaning & Transforming, Encoding (Label Encode) & Imputing (KNN imputation), rectifying classification bias (Imbalance Learn) to develop machine learning model (sci-kit learn, lightGBM) with optimal metrics in Flask framework.
- Creating Sequence diagrams, use case scenarios and performing interface analysis for input validation of model.
- Using feature engineering techniques like Recursive Feature Selection, SelectfromModel, PCA to improvise model metrics.
- Assessing threshold values using Hyper-parameter tuning, k-fold cross validation to configure model parameters for optimal accuracy of the model.
- Performing User Acceptance Testing before production release.

#### Tool Stack:

- Python, My SQL, Flask, Machine Learning (scikit-learn, lightGBM, imbalance learn, pandas, scipy), Figma, MS Visio.

#### Capstone Project (Master's Research Project)

#### Smart Farming IOT Mobile application

#### Swinburne University of Technology

#### Role : Business Technology Analyst

Hawthorn, VIC

08/2020 - 12/2020

**Project Description:** A Flutter application built to monitor farm atmosphere, quality of milk based on IOT data and notify farmers about an anomaly with regards to temperature of milk silo and weather conditions. It also helps to notify farmers about the schedule and location of the milk truck.

#### Responsibilities

- Documentation of Functional requirements, Non-Functional requirements, system requirements and technology stack in Software Requirement Specification document.

- Assisting in Product Features, use cases, Wireframe design and prototyping in Adobe XD.
- Designing system architecture, sequence diagram and data flow diagram in MS Visio.
- Developing navigation drawer and basic layout for the application using interface analysis.
- Developing application code to streamline IOT sensor data through API to notify temperature or atmospheric changes on farm.
- Code cleaning and User acceptance testing before deployment.

**Technology Stack:**

- Adobe XD, Stoplight API, Microsoft Visio, Flutter (Dart), Cross-platform Deployment (iOS , android)