



Onboarding Meeting

[September 8, 2020]



AGENDA

Meeting Duration: 1 Hour (Approx.)

- Introduction to fuse|machines team
- Introduction to Signa Sports team
- Project context
- Brainstorming / Requirements
- Project management practises
- Communication frequency and channel
- Q & A
- Next Steps

TEAM MEMBERS

Data

- Vivek Singh (Data Architect)
- Suman Kumar (Senior Data Lead)
- Tchiring Pemba Lama (Data Engineer)

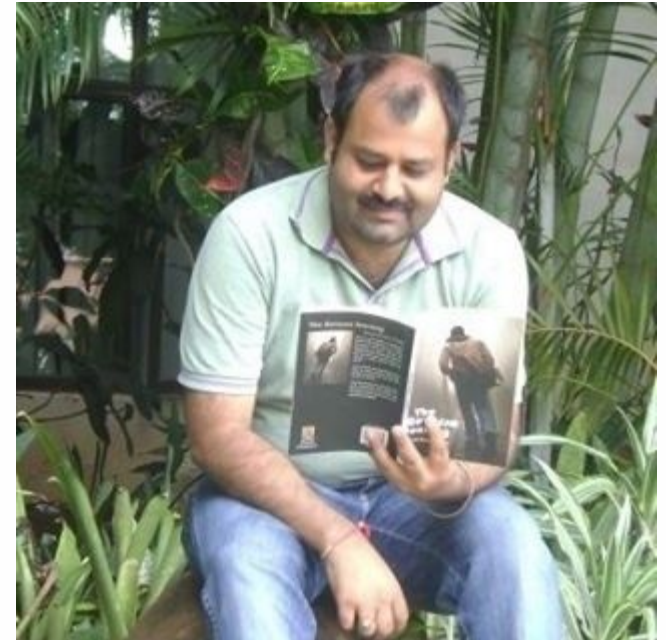
ML

- Bulent Uyaniker PHD (ML Lead)
- Kamal Shrestha (ML Engineer)
- Rojen B.N Pradhan (ML Engineer)

TEAM MEMBERS

Vivek Singh

- Data Architect
- Has unique blend of building high scalable, high availability, data intensive applications which functions in near real time.
- Covered all major blocks of data life cycle - semantic layer, data ingestion, data quality, data processing, data science, master data management, data governance.
- Creator of osDQ. Spark data pipeline. eX-Yahoo! Bigdata Architect



TEAM MEMBERS

Suman Kumar

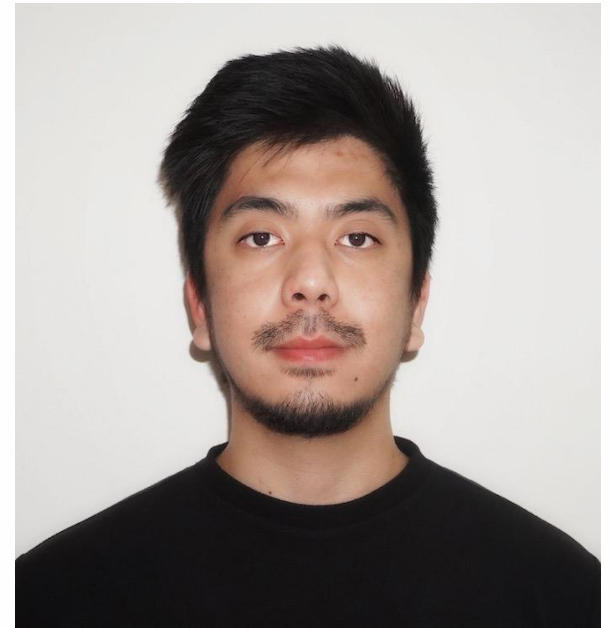
- Senior Data Lead
- 10 years as Big Data Cloud Engineer
- Well versed in AWS, GCP cloud for storage & compute capabilities
- Hands-on with Big Data, Spark, Dataflow, Streaming, Data Lake in On-prem, Cloud
- Design data pipelines from data ingestion to complex processing and retrieval



TEAM MEMBERS

Tchiring Pemba Lama

- Data Engineer
- A seasoned ETL developer with experience in building pipelines with a multitude of big data tools.
- Worked on a variety of projects with different kinds of dataset keeping confidentiality and security in mind.
- Hands on experience with administrative task related to data engineering tools.



TEAM MEMBERS

Bulent Uyaniker, PhD

- ML Lead
- Experienced physicist trained in Data Science
- Skilled in every aspect of Algorithm and Software development, signal processing, statistical analysis and handling large data sets.
- Managed projects, led research and technical groups. Strong skills in building effective teams.



TEAM MEMBERS

Kamal Shrestha

- ML Engineer
- Course Content Developer: ML, DL, CS for AI, Intro-to-AI, Fuse Academics
- Hands-on experience in building customized solutions for deep learning, natural language processing, and machine learning problems on structured data.
- Experience in data wrangling, feature extraction, imputation and selection, generating models, training, evaluation and deployment.



TEAM MEMBERS

Rojen B.N Pradhan

- ML Engineer
- Course Content Developer for Fuse Academics. Created content for international courses like ML, DL, NLP, AI in Biz and, Mathematics.
- Proficiency in building customized machine learning and deep learning solutions for natural language processing, computer vision and regression/classification problems.
- Hands-on experience in implementing complete machine learning pipeline including exploratory data analysis, data wrangling, feature engineering, model generation, training, evaluation and deployment using variety of tools and techniques.
- [Link to CV](#)



SIGNA MEMBERS

- Group CTO - [Thomas Neumann](#)
- Head of Data & BI - [Daniel Lubos](#)
- Data Science - Christian Mena



PROJECT CONTEXT

- Brief background on Signa Sports Client
- Signa Sport's vision



PROJECT REQUIREMENTS

Phase 1: Discovery (Month 1)

Data Engineering

- Conduct Discovery on existing Data Infrastructure and Proposed Architecture
 - Understand the current architecture and data flow
 - Propose the solution, architecture and approach
- Evaluate, design and implement the data pipelines, data lineage and data governance frameworks that are suitable for a modern analytics solution, considering industry standard best practices and patterns.
 - Discovery and assessment
 - Classification and organization
 - Cataloging and metadata management
 - Quality management
 - Access management and governance
 - Compliance
 - Explore benefits of the ETL process being built cloud agnostic
 - Auditing
 - Observability
 - Orchestration
 - Business intelligence/Analytics



PROJECT REQUIREMENTS

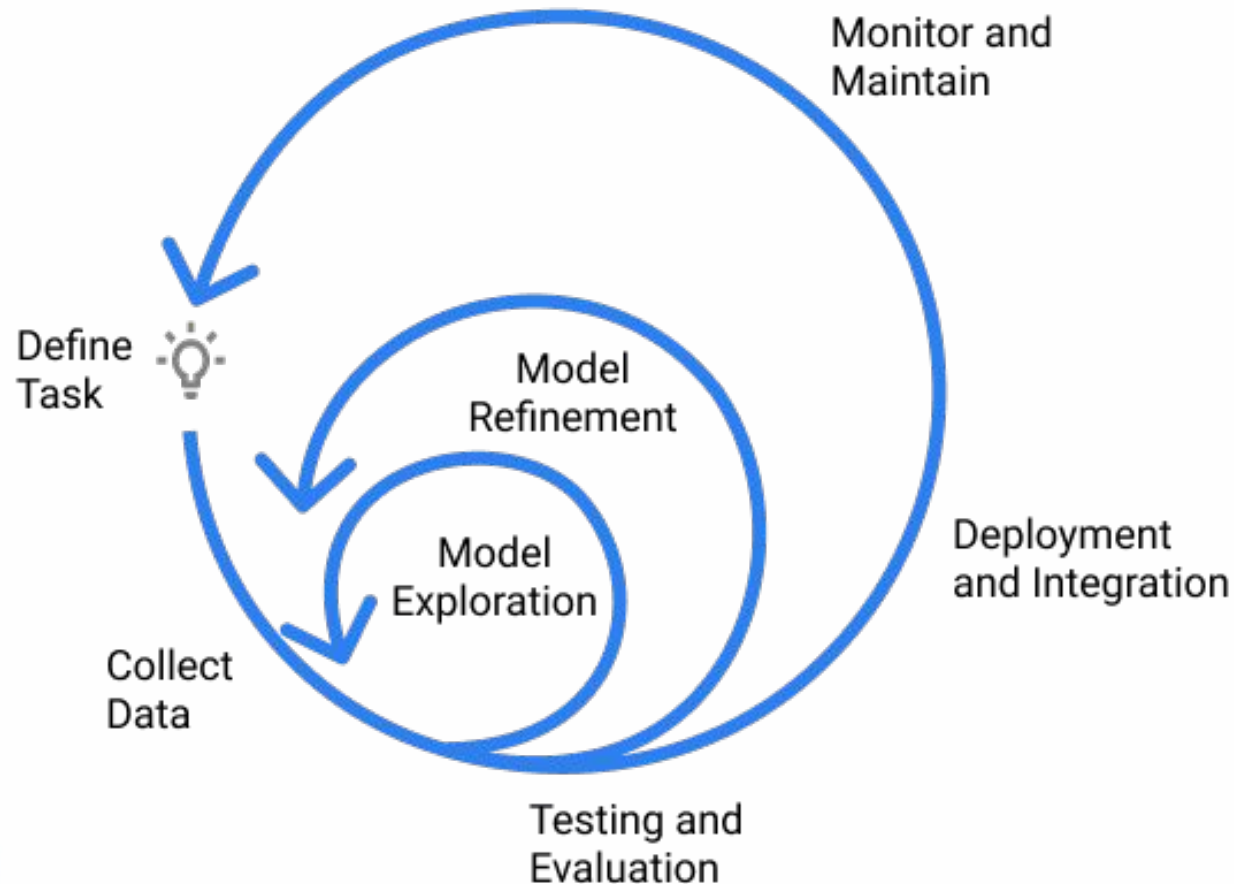
-
- Data migration (if needed)
- Assess best practices and design a schema that matches Signas business needs for delivering a modern analytics solution (descriptive, diagnostic, predictive, prescriptive)
- Collaborate with the Signa Sports United team to deliver a final proposed target architecture and data pipeline solution with a clear SOW for implementation in phase 2.

Data Science/Machine Learning (Price Elasticity)

- Evaluate current model architecture
 - Understand the requirements
 - Assess features and determine additional possible features
 - Run experiments
 - Provide feedback and suggestions

AI SDLC

Machine Learning Development Lifecycle



PROJECT MANAGEMENT PRACTICES

- Development Methodology:
 - Agile and Lean Software Development
- Framework we use:
 - SCRUM
 - Kanban
 - Xtreme Programming



PROJECT MANAGEMENT TOOL

- Project management tools we use
 - Jira



COMMUNICATION

- Channels
 - Slack, Hangouts, **Teams**, etc.
- ~~Frequency~~
 - ~~Daily standup~~
 - ~~Weekly debrief~~



Emails

Data Team

- **Vivek Singh** (Sr Data Architect) - vivek@fusemachines.com
- **Suman Kumar** (Lead Data Engineer) - suman.kumar@fusemachines.com
- **Tchiring Pemba Lama** (Data Engineer) - tchiring@fusemachines.com

ML Team

- **Bulent Uyaniker** (Machine Learning Phd) - bulent@fusemachines.com
- **Kamal Shrestha** (Machine Learning Engineer) - kamal@fusemachines.com
- **Rojen B.N Pradhan** (Machine Learning Engineer) - rojen@fusemachines.com

Timing

North America East	North America West	Nepal	Berlin
8:00 AM	5:00 AM	5:45 PM	2:00 PM
9:00 AM	6:00 AM	7:45 PM	3:00 PM
10:00 AM	7:00 AM	8:45 PM	4:00 PM
11:00 AM	8:00 AM	9:45 PM	5:00 PM
12:00 PM	9:00 AM	10:45 PM	6:00 PM
1:00 PM	10:00 AM		7:00 PM
2:00 PM	11:00 AM		8:00 PM

DE - Tues 10 am est, ML - Mon 10 am est --- > 1 hr sessions



Notes



THANK YOU

