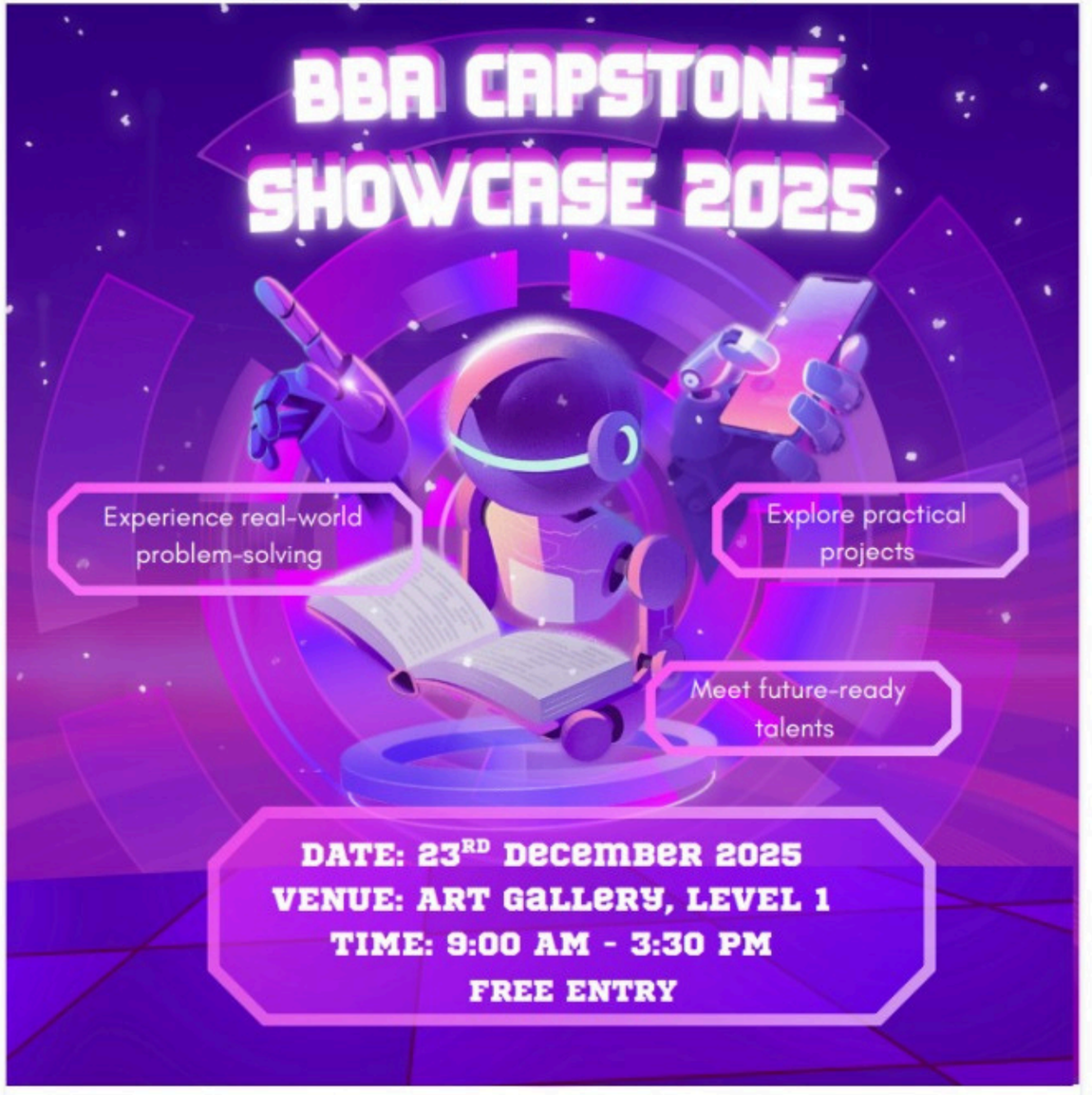
SUNWAY UNIVERSITY
A CLASS ABOVE

SUNWAY COLLEGE
A Brighter Future

BBA CAPSTONE SHOWCASE 2025



DATE: 23RD DECEMBER 2025
VENUE: ART GALLERY, LEVEL 1
TIME: 9:00 AM - 3:30 PM
FREE ENTRY

Sunway Business Analytics

FRIENDS OF LIFE

4 QUALITY EDUCATION

Sunway BUSINESS SCHOOL

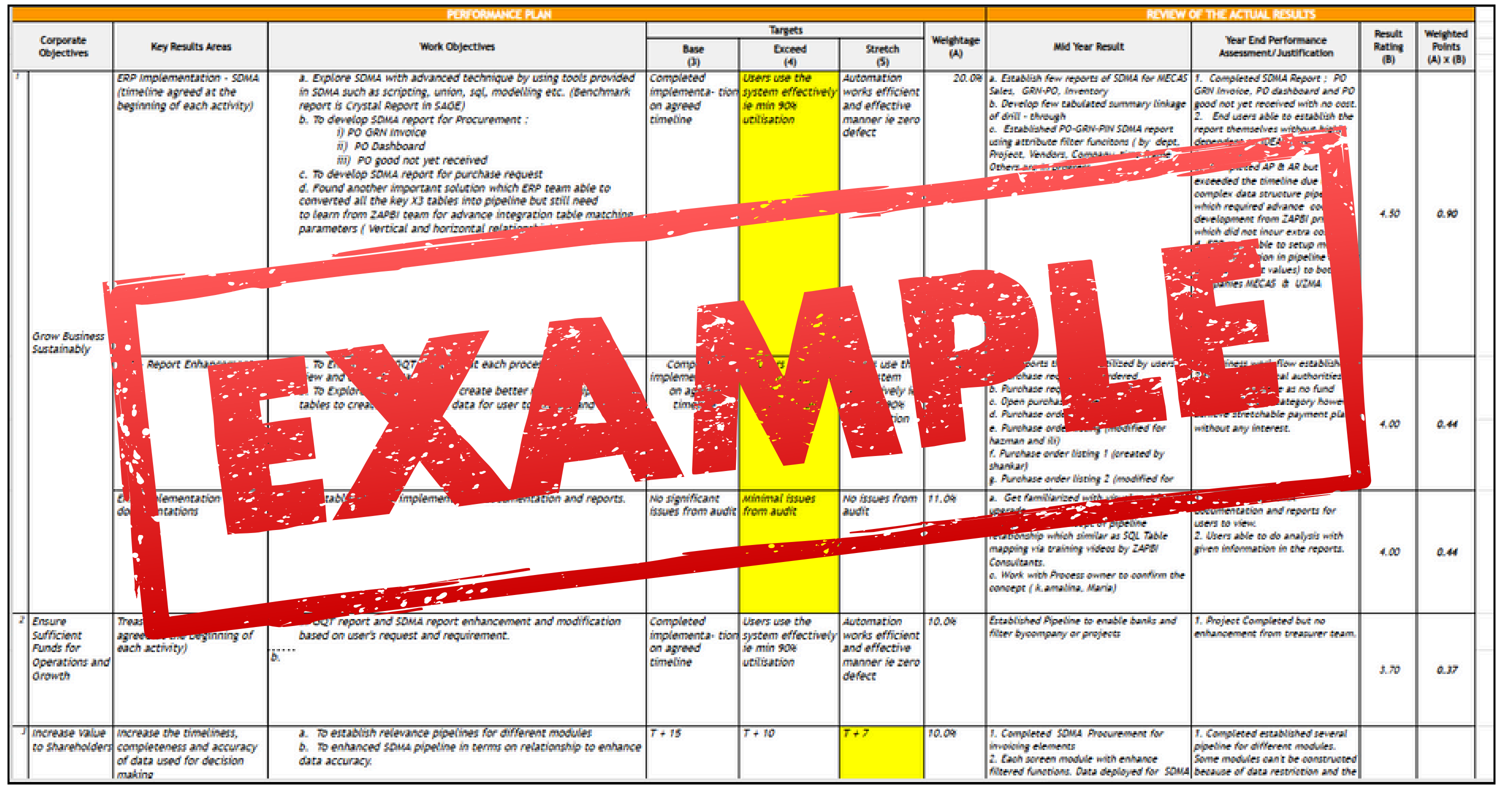
Owned and governed by the
Jeffrey Cheah Foundation

SmartAppraise AI: Predicting Employee Performance Using SMART Data and Non-Linear Feature Analysis

Members: Yaw Chin Hen 21058417, Tam Sze Yin 21093893, Yang Ka Weng 21095872

What is the problem?

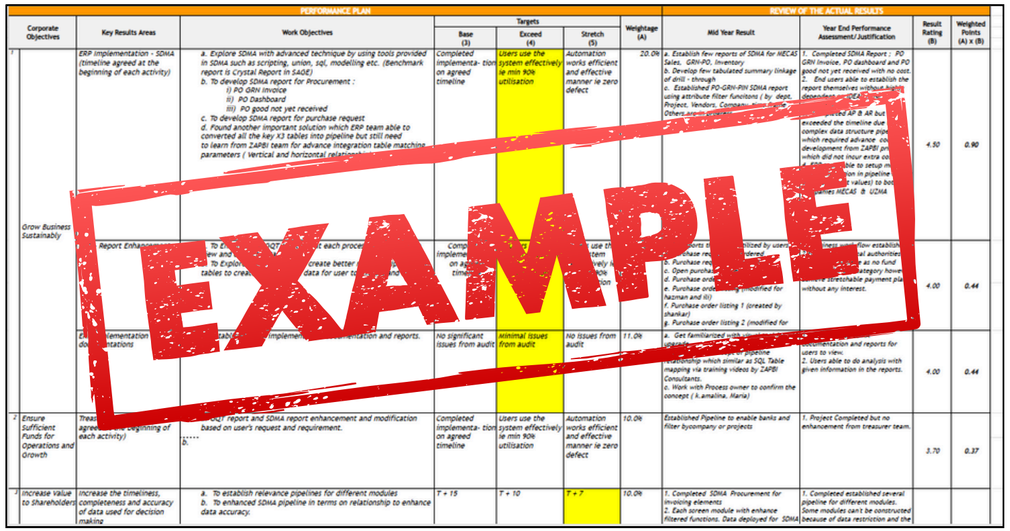
What is the problem?



	PERFORMANCE PLAN			Targets				REVIEW OF THE ACTUAL RESULTS			
	Corporate Objectives	Key Results Areas	Work Objectives	Base (3)	Exceed (4)	Stretch (5)	Weightage (A)	Mid Year Result	Year End Performance Assessment/ Justification	Result Rating (B)	Weighted Points (A) x (B)
1	Grow Business Sustainably	ERP Implementation - SDMA (timeline agreed at the beginning of each activity)	a. Explore SDMA with advanced technique by using tools provided in SDMA such as scripting, union, sql, modelling etc. (Benchmark report is Crystal Report in SAGE) b. To develop SDMA report for Procurement : i) PO GRN Invoice ii) PO Dashboard iii) PO good not yet received c. To develop SDMA report for purchase request d. Found another important solution which ERP team able to converted all the key X3 tables into pipeline but still need to learn from ZAPBI team for advance integration table matching parameters (Vertical and horizontal relations[illegible]	Completed implementa- tion on agreed timeline	Users use the system effectively ie min 90% utilisation	Automation works efficient and effective manner ie zero defect	20.0%	a. Establish few reports of SDMA for MECAS Sales, GRN-PO, Inventory b. Develop few tabulated summary linkage of drill - through c. Established PO-GRN-PIN SDMA report using attribute filter funcitons (by dept, Project, Vendors, Company [illegible] Others [illegible]	1. Completed SDMA Report : PO GRN Invoice, PO dashboard and PO good not yet received with no cost. 2. End users able to establish the report themselves without [illegible] dependen[illegible] [illegible]pleted AP & AR but exceeded the timeline due complex data structure pipe[illegible] which required advance co[illegible] development from ZAPBI pr[illegible] which did not incur extra co[illegible] [illegible]ble to setup m[illegible] [illegible]on in pipeline [illegible] values) to bot[illegible] [illegible]panies MECAS & UZMA	4.50	0.90
		[illegible] Report Enha[illegible]	[illegible] To E[illegible] [illegible]OQT[illegible] [illegible]t each proce[illegible] [illegible]iew and [illegible] To Explor[illegible] [illegible] create better [illegible] tables to crea[illegible] data for user t[illegible] and [illegible]	Comp[illegible] impleme[illegible] on ag[illegible] time[illegible]	[illegible]	[illegible] use th[illegible] [illegible]tem [illegible]vely i[illegible] [illegible] 90% [illegible]ion	[illegible]	[illegible]ports t[illegible] [illegible]lized by users [illegible]rchase req[illegible] [illegible]rdered b. Purchase req[illegible] c. Open purchas[illegible] d. Purchase ord[illegible] e. Purchase orde[illegible] (modified for hazman and ili) f. Purchase order listing 1 (created by shankar) g. Purchase order listing 2 (modified for	[illegible]iness [illegible] flow establish[illegible] [illegible]al authorities [illegible] as no fund [illegible]ategory howe[illegible] [illegible] stretchable payment pla[illegible] without any interest.	4.00	0.44
		E[illegible]lementation do[illegible]ntations	[illegible]tabl[illegible] implemen[illegible] [illegible]entation and reports.	No significant issues from audit	Minimal issues from audit	No issues from audit	11.0%	a. Get familiarized with [illegible] upgrade [illegible] [illegible] of pipeline relationship which similar as SQL Table mapping via training videos by ZAPBI Consultants. c. Work with Process owner to confirm the concept (k.amalina, Maria)	[illegible] documentation and reports for users to view. 2. Users able to do analysis with given information in the reports.	4.00	0.44
2	Ensure Sufficient Funds for Operations and Growth	Treas[illegible] agree[illegible] beginning of each activity)	[illegible]OQT report and SDMA report enhancement and modification based on user's request and requirement. b.	Completed implementa- tion on agreed timeline	Users use the system effectively ie min 90% utilisation	Automation works efficient and effective manner ie zero defect	10.0%	Established Pipeline to enable banks and filter bycompany or projects	1. Project Completed but no enhancement from treasurer team.	3.70	0.37
3	Increase Value to Shareholders	Increase the timeliness, completeness and accuracy of data used for decision making	a. To establish relevance pipelines for different modules b. To enhanced SDMA pipeline in terms on relationship to enhance data accuracy.	T + 15	T + 10	T + 7	10.0%	1. Completed SDMA Procurement for invoicing elements 2. Each screen module with enhance filtered functions. Data deployed for SDMA	1. Completed established several pipeline for different modules. Some modules can't be constructed because of data restriction and the		

SMART KPI was introduced to overcome biases but...

What is the problem?



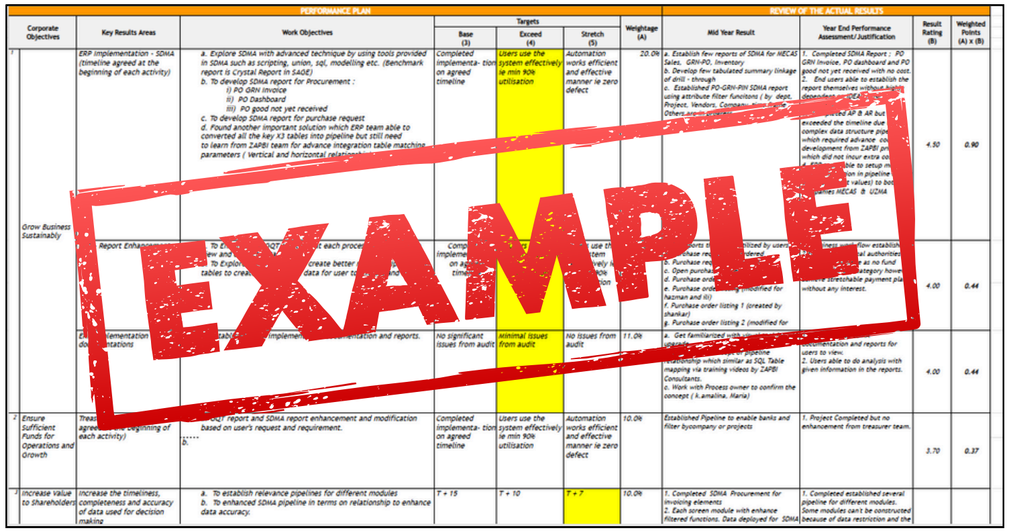
Corporate Objectives	Key Results Areas	Work Objectives	Targets: Base (3)	Targets: Exceed (4)	Targets: Stretch (5)	Weightage (A)	Mid Year Result	Year End Performance Assessment/ Justification	Result Rating (B)	Weighted Points (A) x (B)
PERFORMANCE PLAN							REVIEW OF THE ACTUAL RESULTS			
1 Grow Business Sustainably	ERP implementation - SDHA (timeline agreed at the beginning of each activity)	a. Explore SDHA with advanced technique by using tools provided in SDHA such as scripting, union, sql, modelling etc. (Benchmark report is Crystal Report in SAGE) b. To develop SDHA report for Procurement : i) PO GRN invoice ii) PO Dashboard iii) PO good not yet received c. To develop SDHA report for purchase request d. Found another important solution which ERP team able to converted all the key X3 tables into pipeline but still need to learn from ZAPBI team for advance integration table matching parameters (Vertical and horizontal relationship)	Completed implementa- tion on agreed timeline	Users use the system effectively ie min 90% utilisation	Automation works efficient and effective manner ie zero defect	20.0%	a. Establish few reports of SDHA for HECAS Sales, GRN-PO, Inventory b. Develop few tabulated summary linkage of drill - through c. Established PO-GRN-PIN SDHA report using attribute filter funcitons (by dept, Project, Vendors, Compa... Others...are in production)	1. Completed SDHA Report : PO GRN Invoice, PO dashboard and PO good not yet received with no cost. 2. End users able to establish the report themselves without highly dependent HECAS ...ified AP & AR but exceeded the timeline due complex data structure pipe... which required advance co... development from ZAPBI pr... which did not incur extra co... ...able to setup m... ...tion in pipeline ...r values) to bu... ...ities HECAS & UCHA	4.50	0.90
	Report Enhan...	To E... ...ew and ... To Explo... ...tables to crea...	Comp... Implementa... on ag... time...	...use th... ...ively ...ble ...tion			...erts t... ...hase re... c. Open purcha... d. Purchase orde... e. Purchase orde... ...modified for hazman and iii) f. Purchase order listing 1 (created by shankar) g. Purchase order listing 2 (modified for	...ilized by users ...dered ... ...flow establish... ...al authorities ... ...t as no fund ...ategory hover... ...stretchable payment pla... without any interest.	4.00	0.44
	Tr... Implementation ...documentations	To establ... ...Implement... ...mentation and reports.	No significant issues from audit	Minimal issues from audit	No issues from audit	11.0%	a. Get familiarized with ... ...operations ... ...pipeline. Relationship which similar as SQL Table mapping via training videos by ZAPBI Consultants. c. Work with Process owner to confirm the concept (k.amalina, Haria)	...documentation and reports for users to view. 2. Users able to do analysis with given information in the reports.	4.00	0.44
2 Ensure Sufficient Funds for Operations and Growth	Trea... agre... ...beginning of each activity)	a. To develop ...OGY report and SDHA report enhancement and modification based on user's request and requirement. b.	Completed implementa- tion on agreed timeline	Users use the system effectively ie min 90% utilisation	Automation works efficient and effective manner ie zero defect	10.0%	Established Pipeline to enable banks and filter bycompany or projects	1. Project Completed but no enhancement from treasurer team.	3.70	0.37
3 Increase Value to Shareholders	Increase the timeliness, completeness and accuracy of data used for decision making	a. To establish relevance pipelines for different modules b. To enhanced SDHA pipeline in terms on relationship to enhance data accuracy.	T + 15	T + 10	T + 7	10.0%	1. Completed SDHA Procurement for invoicing elements 2. Each screen module with enhance filtered functions. Data deployed for SDHA	1. Completed established several pipeline for different modules. Some modules can't be constructed because of data restriction and the		

EXAMPLE

SMART KPI was introduced to overcome biases but...

The Hidden Challenges of SMART KPI

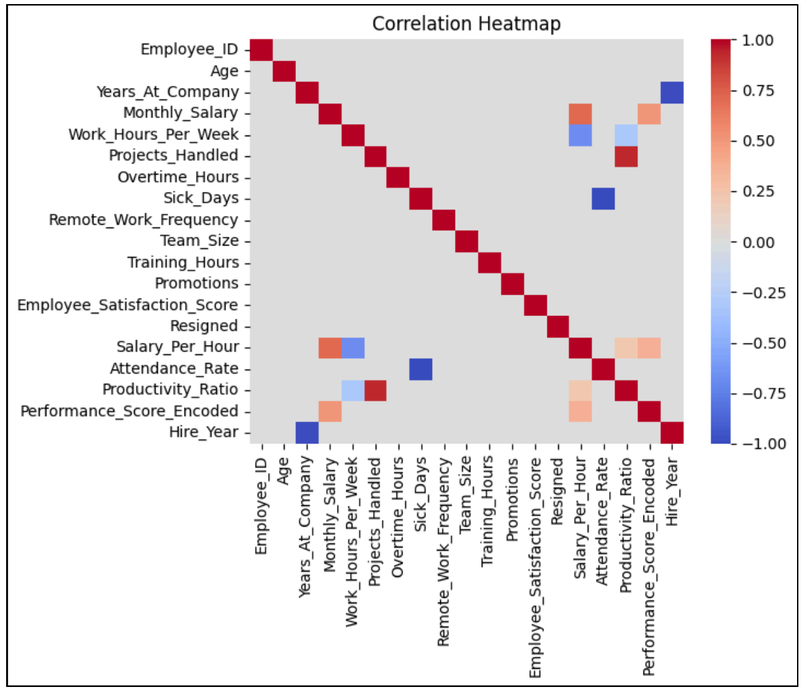
What is the problem?



SMART KPI was introduced to overcome biases but...

The Hidden Challenge of SMART KPI

- KPI data is non-linear
- KPIs are often weakly correlated
- Manual analysis becomes complex at scale

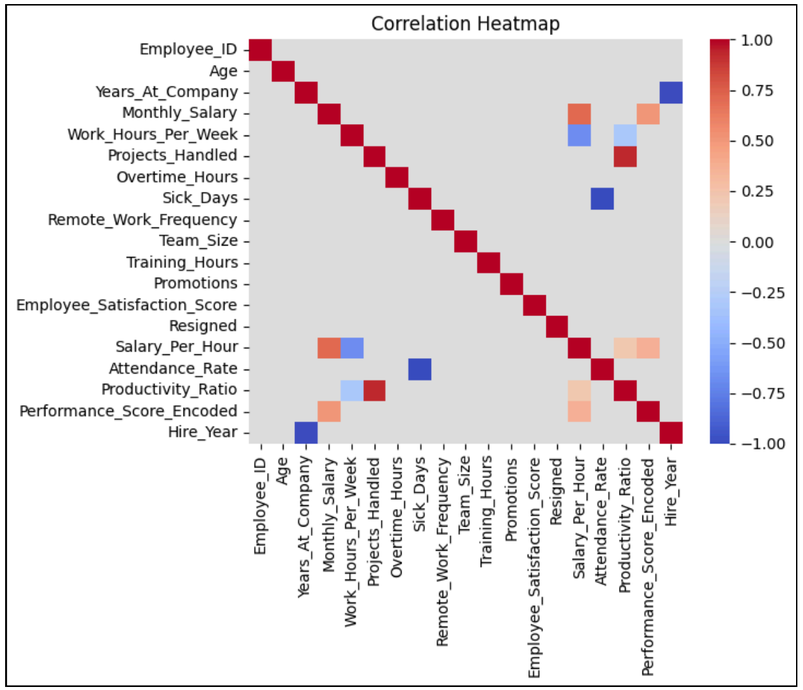


What is the problem?

The Hidden Challenge of SMART KPI

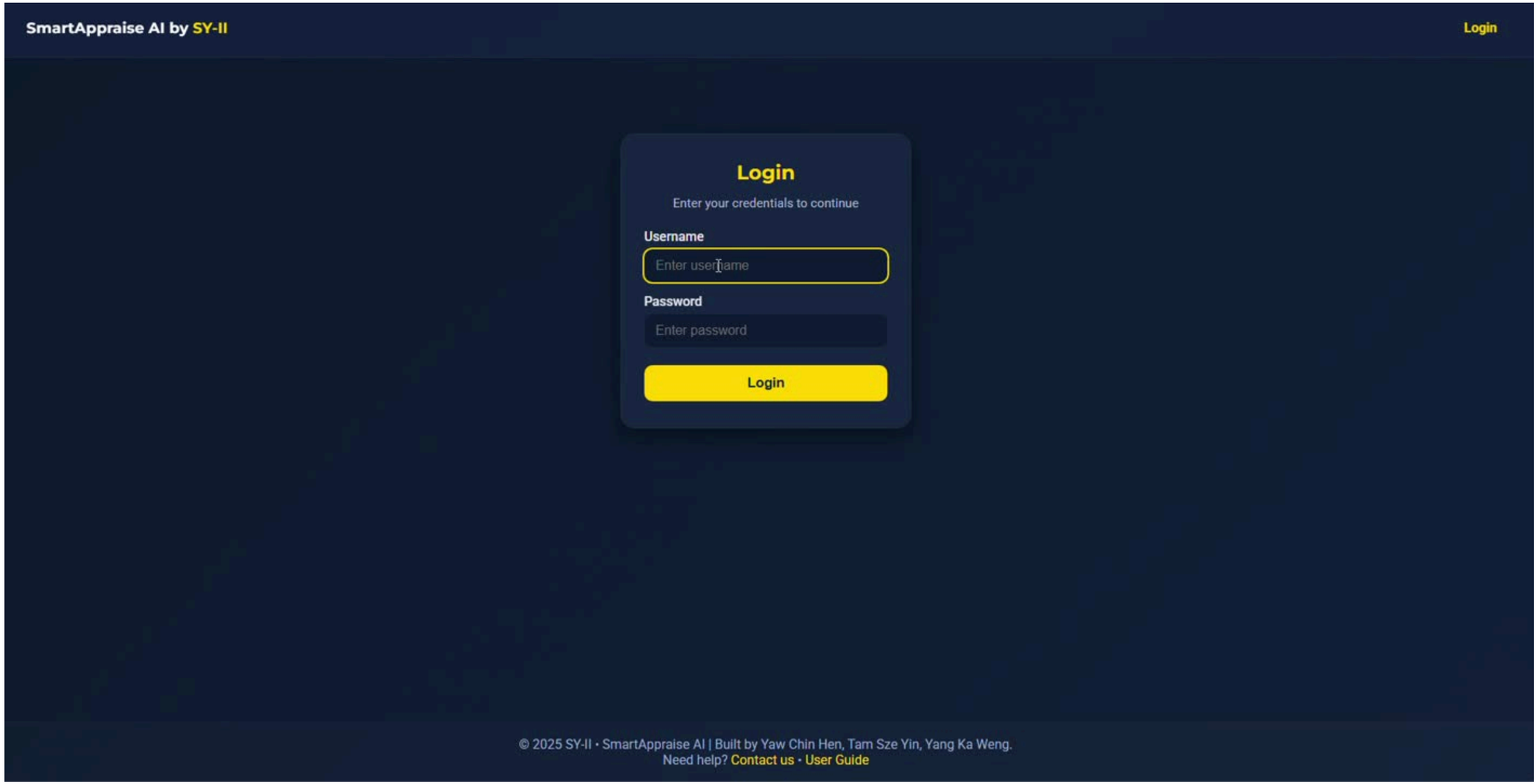
- KPI data is non-linear
- KPIs are often weakly correlated
- Manual analysis becomes complex at scale

Fair
Consistent
Data-driven

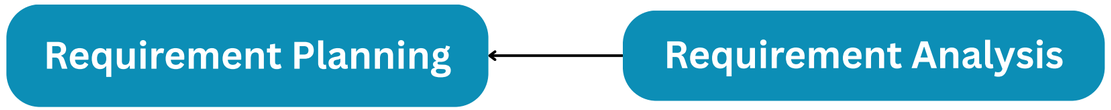


Objective and Solution

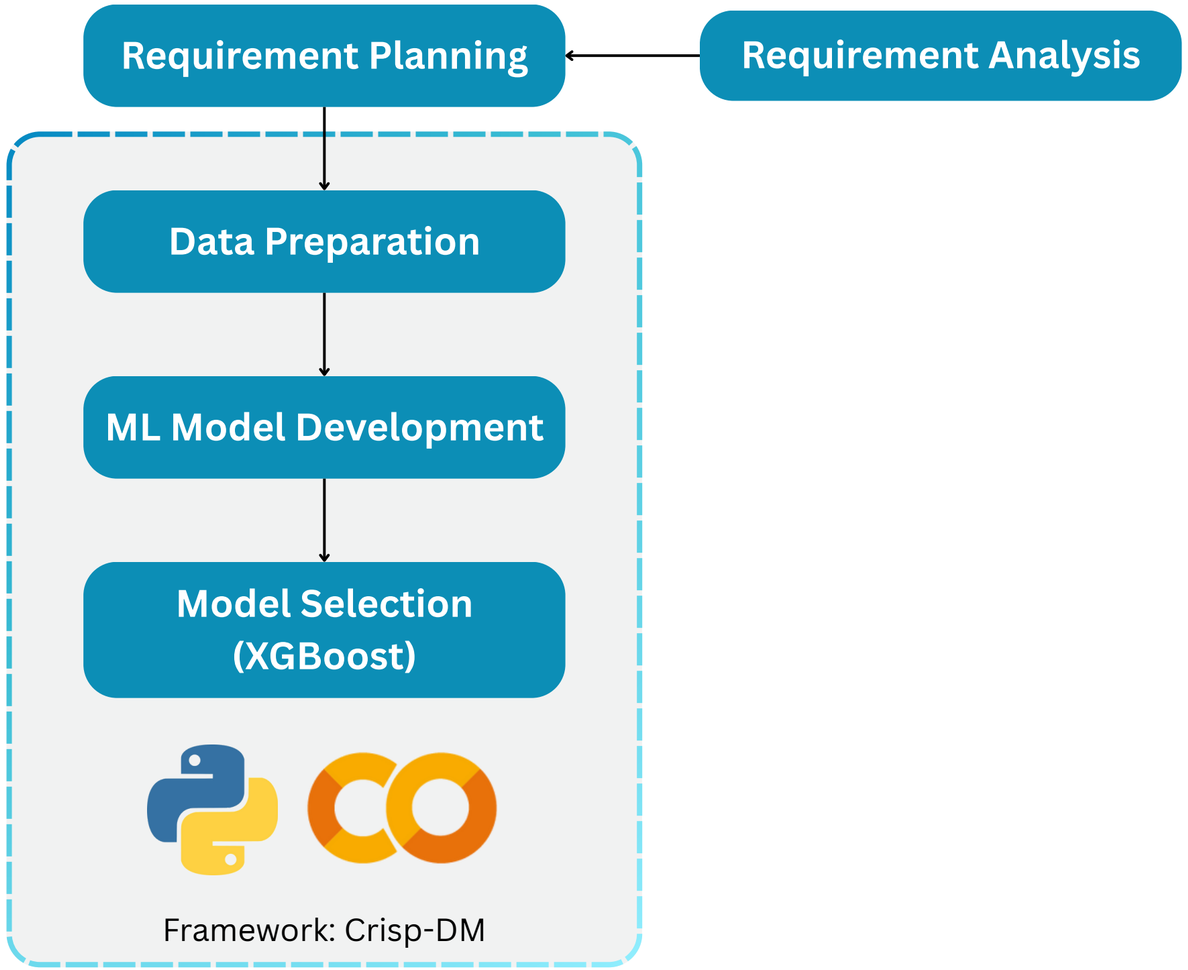
Performance appraisal platform that leverages **SMART KPI data and non-linear feature analysis** to deliver **objective, data-driven, and bias-reduced** employee performance predictions.



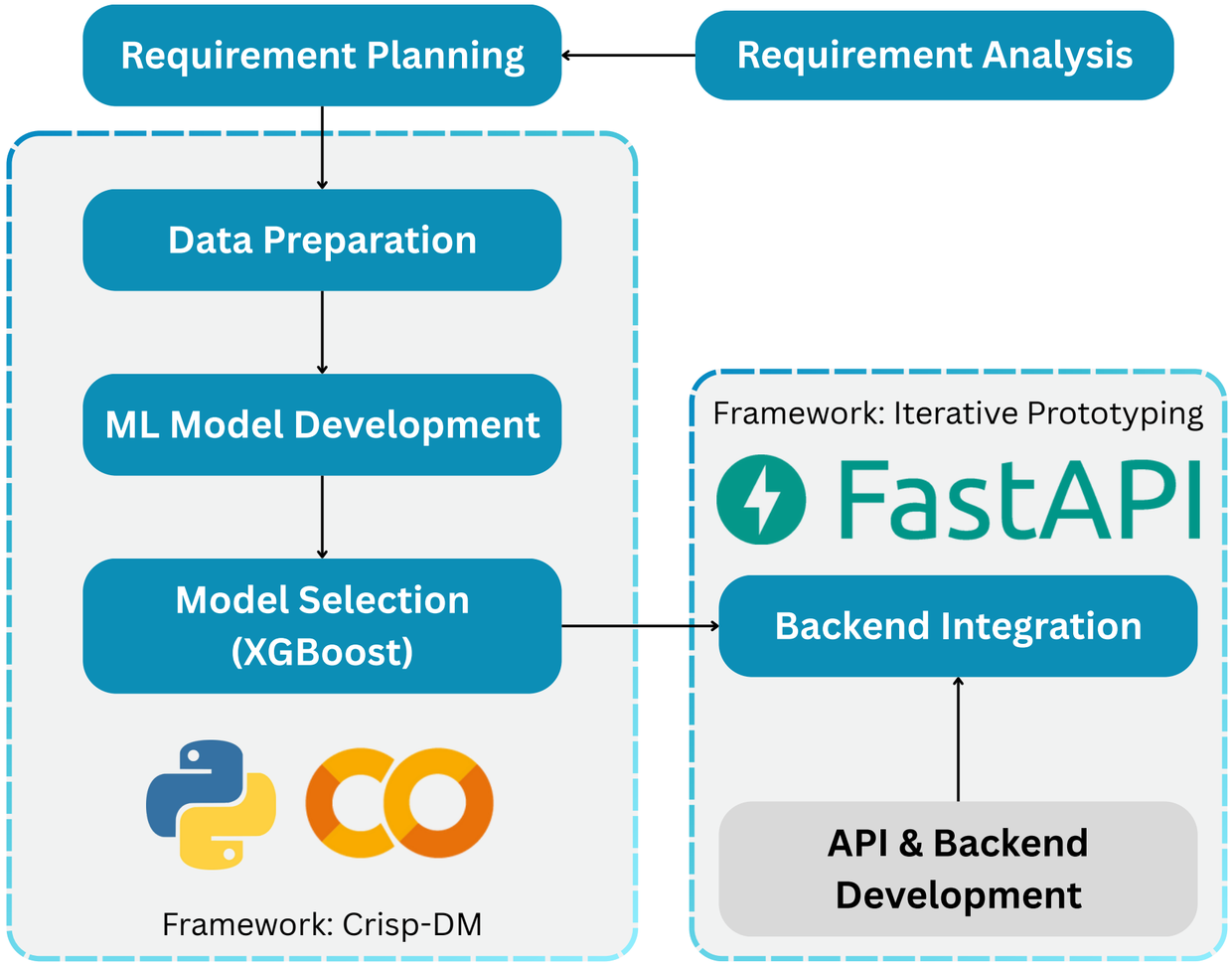
Project Life Cycle



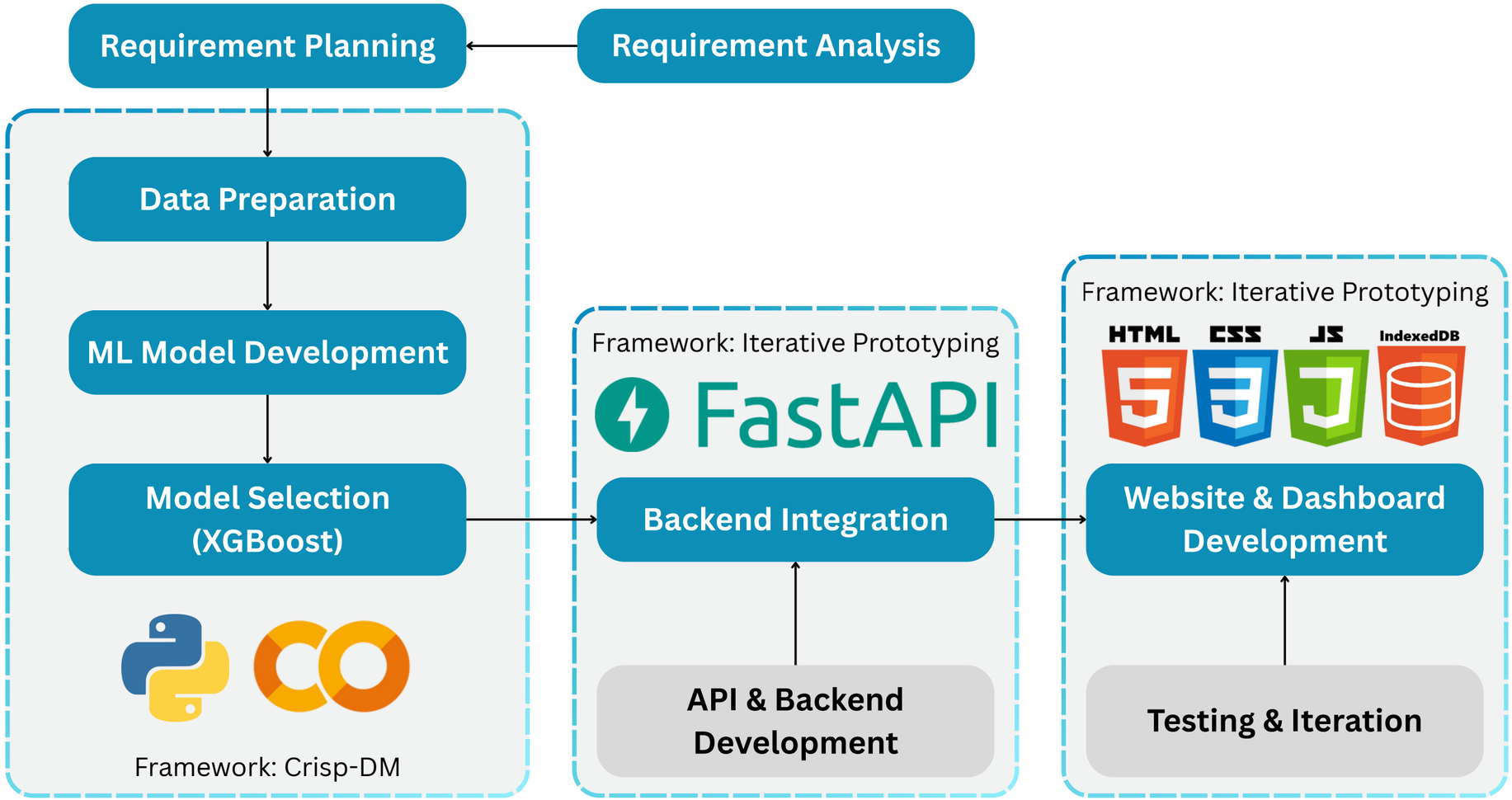
Project Life Cycle



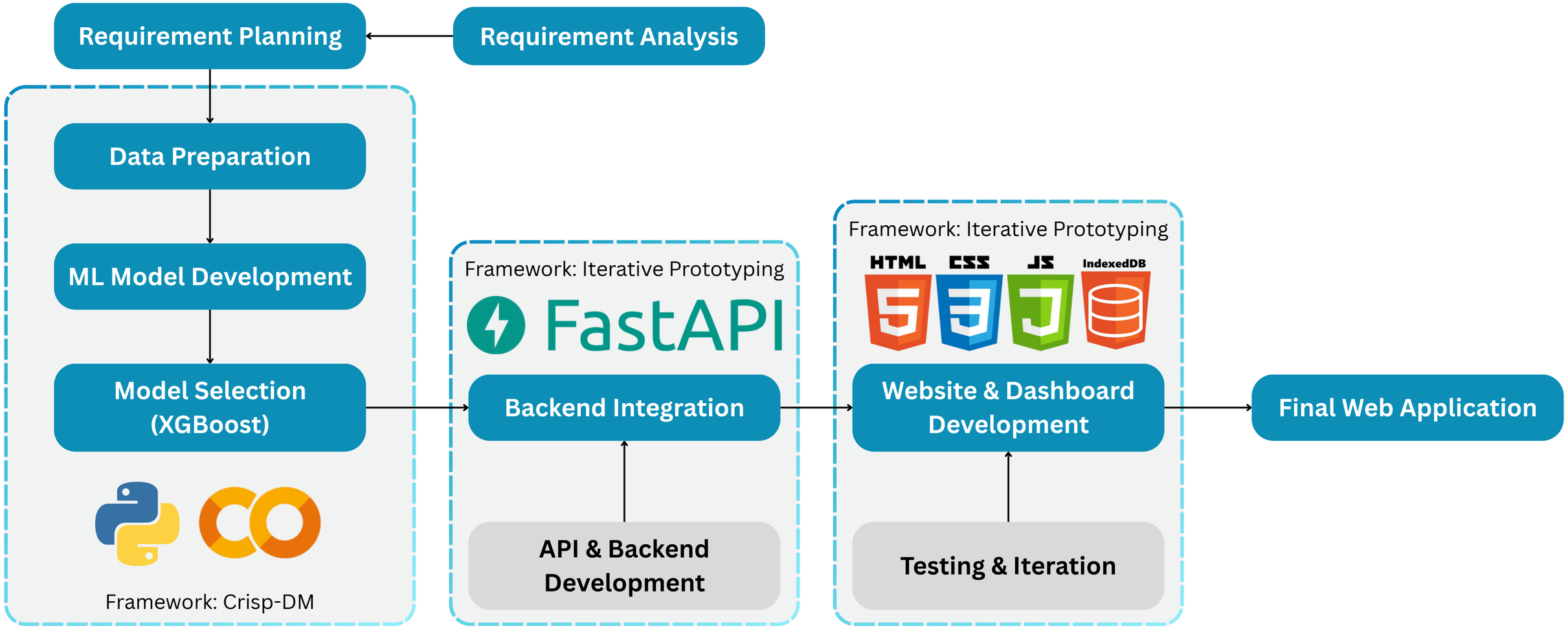
Project Life Cycle



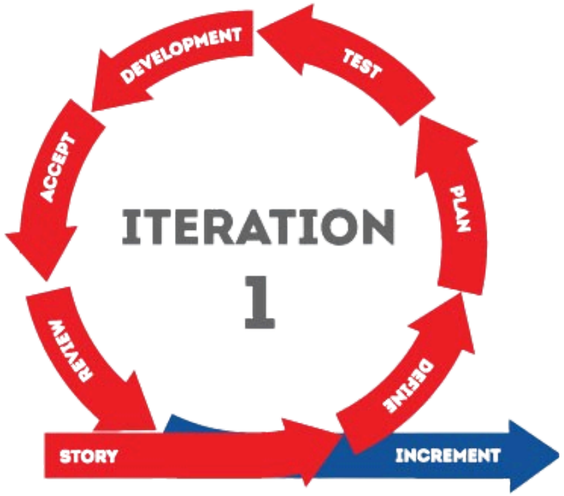
Project Life Cycle



Project Life Cycle



Web Development Life Cycle

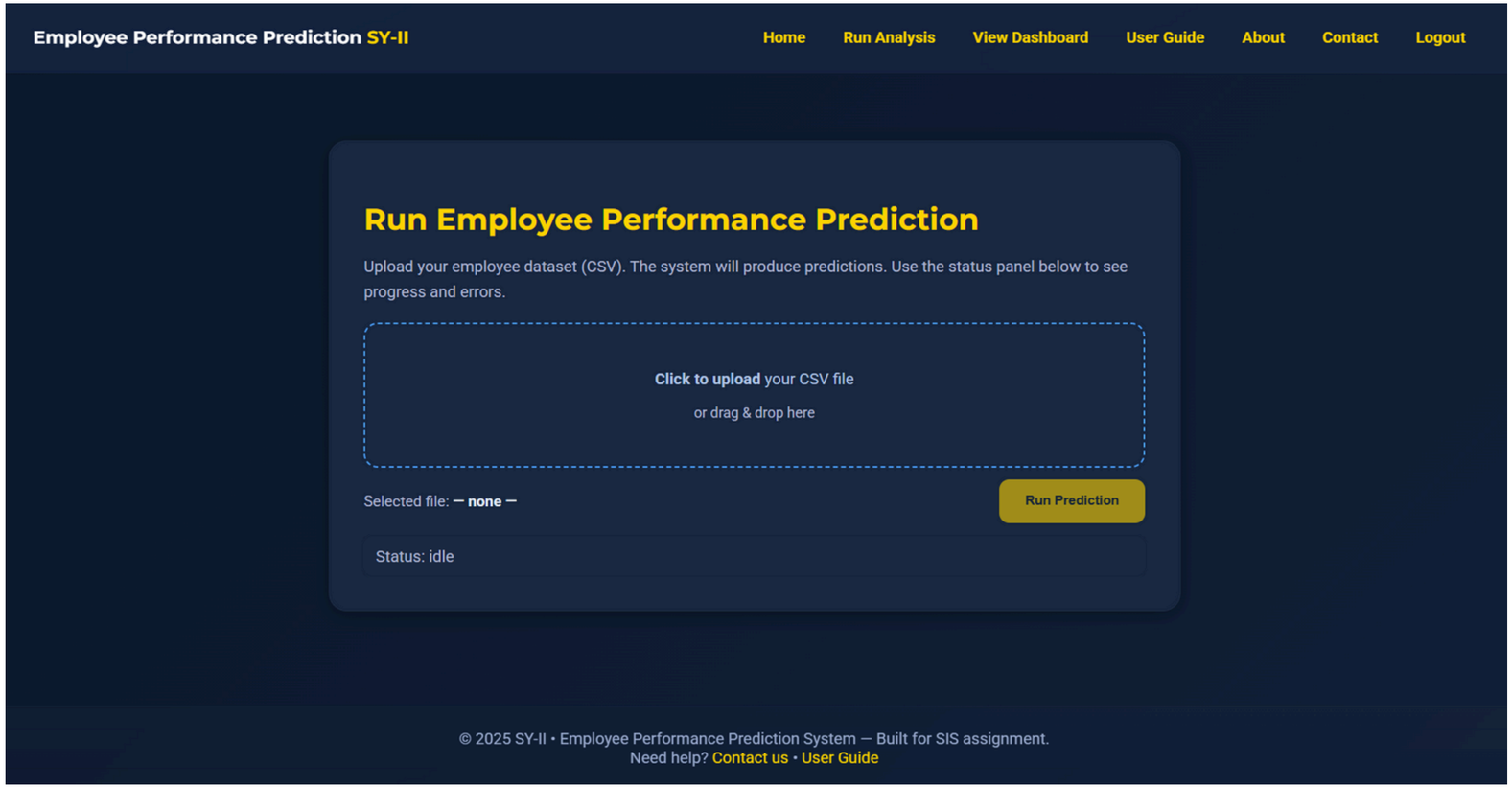


Iteration 1: File upload interface + backend API integration

Focus: Backend Logic & Data Ingestion

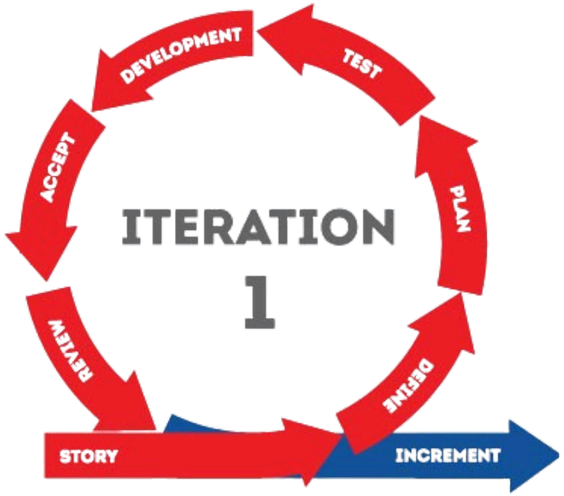
Deliverables: File Upload Interface. Backend API integration with XGBoost Model. Handling data preprocessing pipeline (cleaning/formatting).

Technical Contribution: Established the communication bridge between the web server and the ML engine.



Source: (Hulugh & Emmanuel, 2024)

Web Development Life Cycle

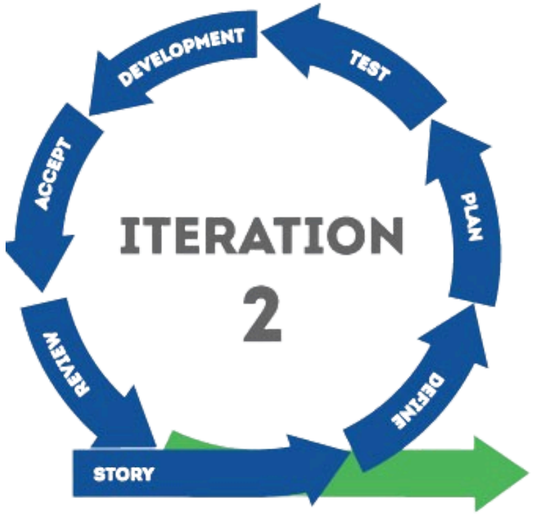


Iteration 1: File upload interface + backend API integration

Focus: Backend Logic & Data Ingestion

Deliverables: File Upload Interface. Backend API integration with XGBoost Model. Handling data preprocessing pipeline (cleaning/formatting).

Technical Contribution: Established the communication bridge between the web server and the ML engine.

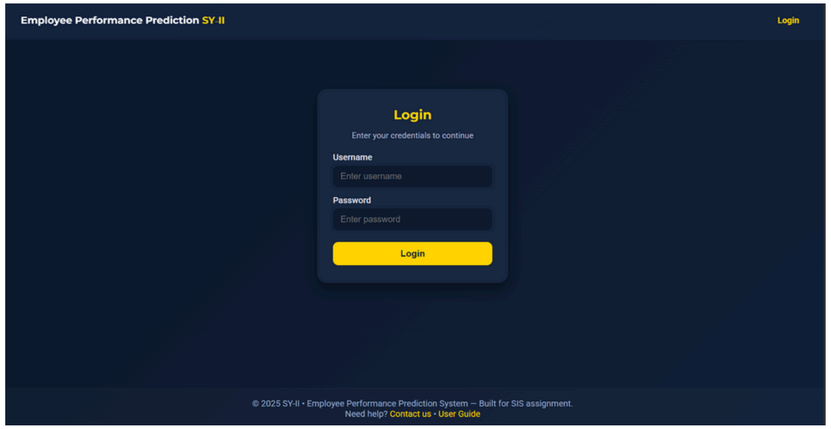


Iteration 2: Dashboard development & visualization

Focus: Dashboarding & Results

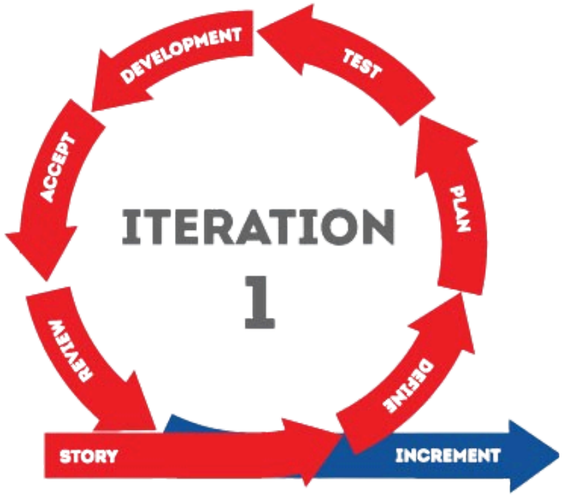
Deliverables: Real-time charts and prediction tables. Data aggregation for visual insights.

Technical Contribution: Transformed raw JSON API responses into user-friendly graphical interfaces.



Source: (Hulugh & Emmanuel, 2024)

Web Development Life Cycle

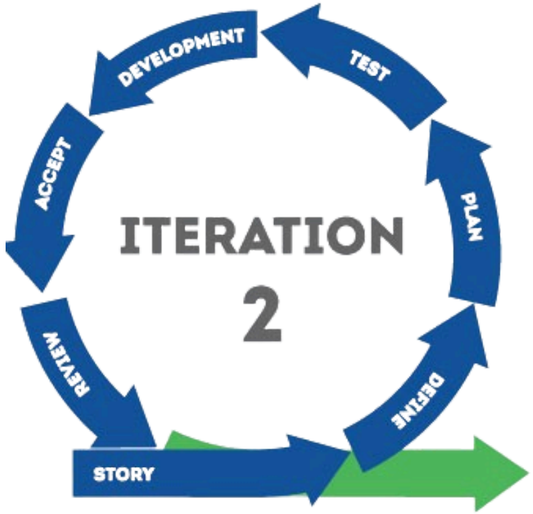


Iteration 1: File upload interface + backend API integration

Focus: Backend Logic & Data Ingestion

Deliverables: File Upload Interface. Backend API integration with XGBoost Model. Handling data preprocessing pipeline (cleaning/formatting).

Technical Contribution: Established the communication bridge between the web server and the ML engine.



Iteration 2: Dashboard development & visualization

Focus: Dashboarding & Results

Deliverables: Real-time charts and prediction tables. Data aggregation for visual insights.

Technical Contribution: Transformed raw JSON API responses into user-friendly graphical interfaces.

Iteration 3: Refinement

Focus: Stability & UI/UX

Deliverables: Conducting End-to-End System Testing & Bug Fixes. UI Polish. Final Validation of prediction accuracy.

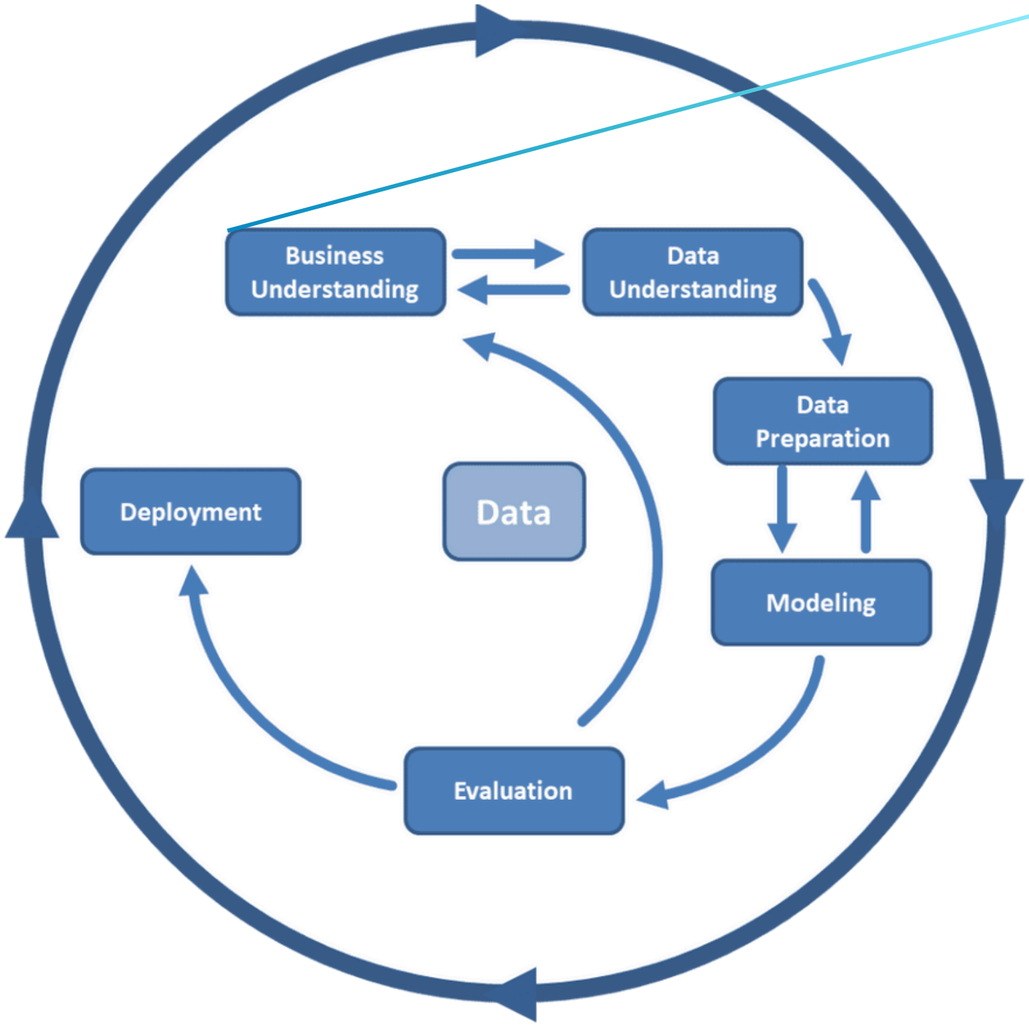
Technical Contribution: Ensured system reliability and finalized the user experience for deployment.

Source: (Hulugh & Emmanuel, 2024)

ML Project Life Cycle

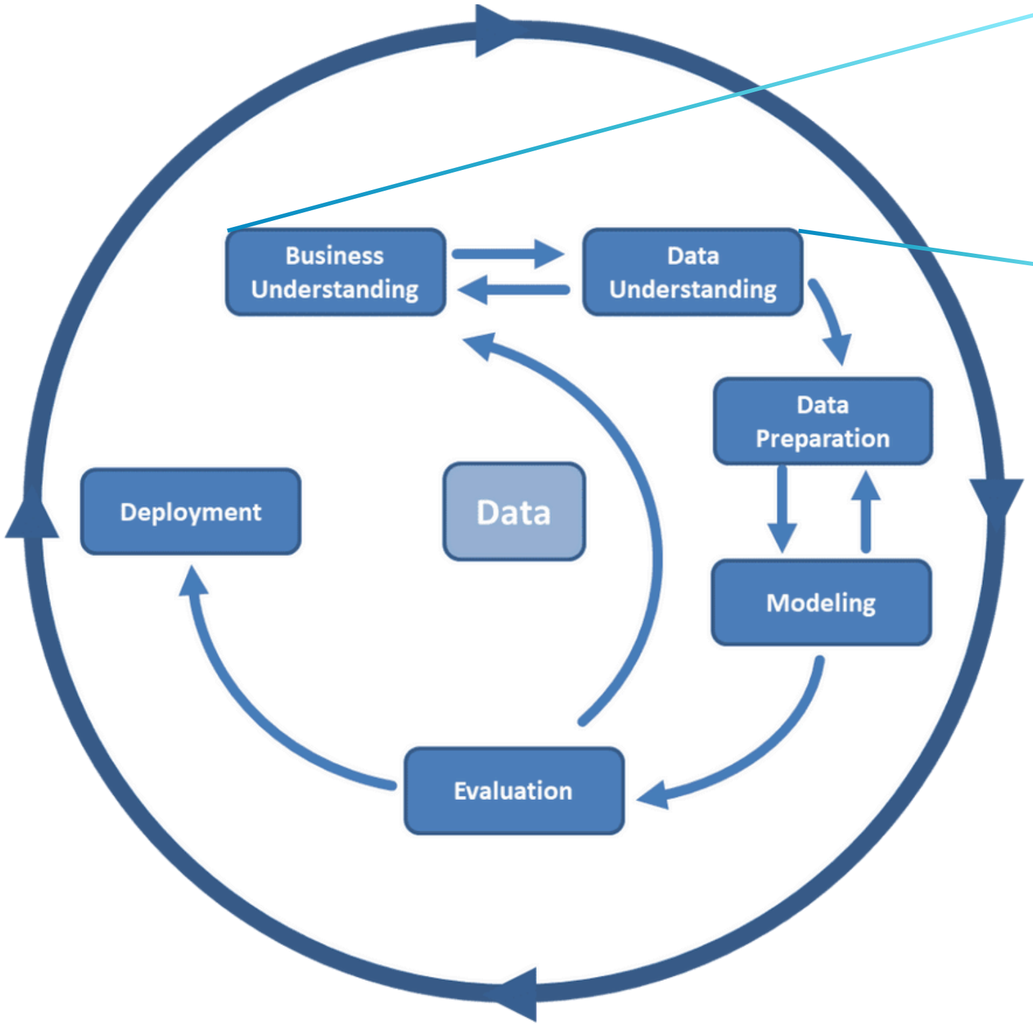
1. Business Understanding

- To predict employee performance class using SMART KPI.



Source: (Muhammad Fawwaz Narendra, 2025)

ML Project Life Cycle



1. Business Understanding

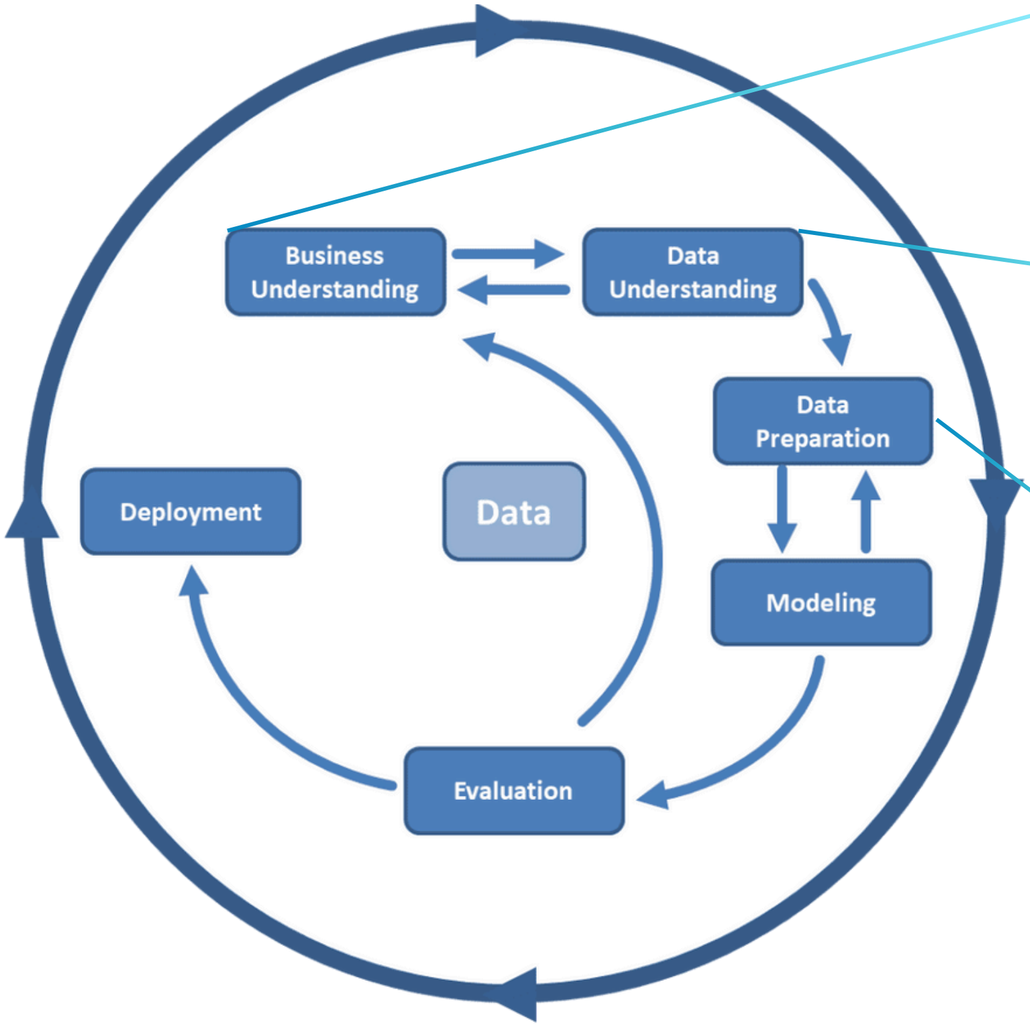
- To predict employee performance class using SMART KPI.

2. Data Understanding

- Explore dataset structure, feature, distributions, and anomalies.

Source: (Muhammad Fawwaz Narendra, 2025)

ML Project Life Cycle



1. Business Understanding

- To predict employee performance class using SMART KPI.

2. Data Understanding

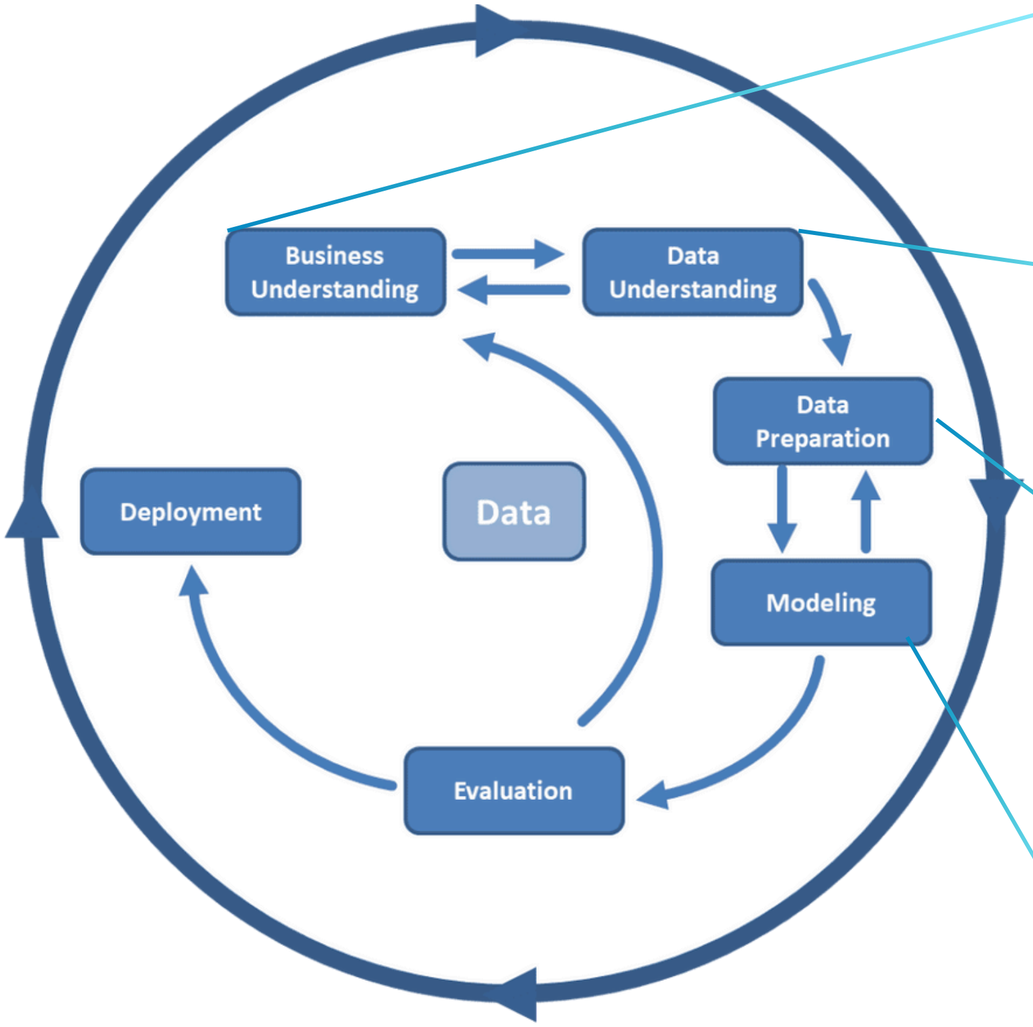
- Explore dataset structure, feature, distributions, and anomalies.

3. Data Preparation

- Data cleaning
- Feature engineering
- Label encoding for performance categories

Source: (Muhammad Fawwaz Narendra, 2025)

ML Project Life Cycle



1. Business Understanding

- To predict employee performance class using SMART KPI.

2. Data Understanding

- Explore dataset structure, feature, distributions, and anomalies.

3. Data Preparation

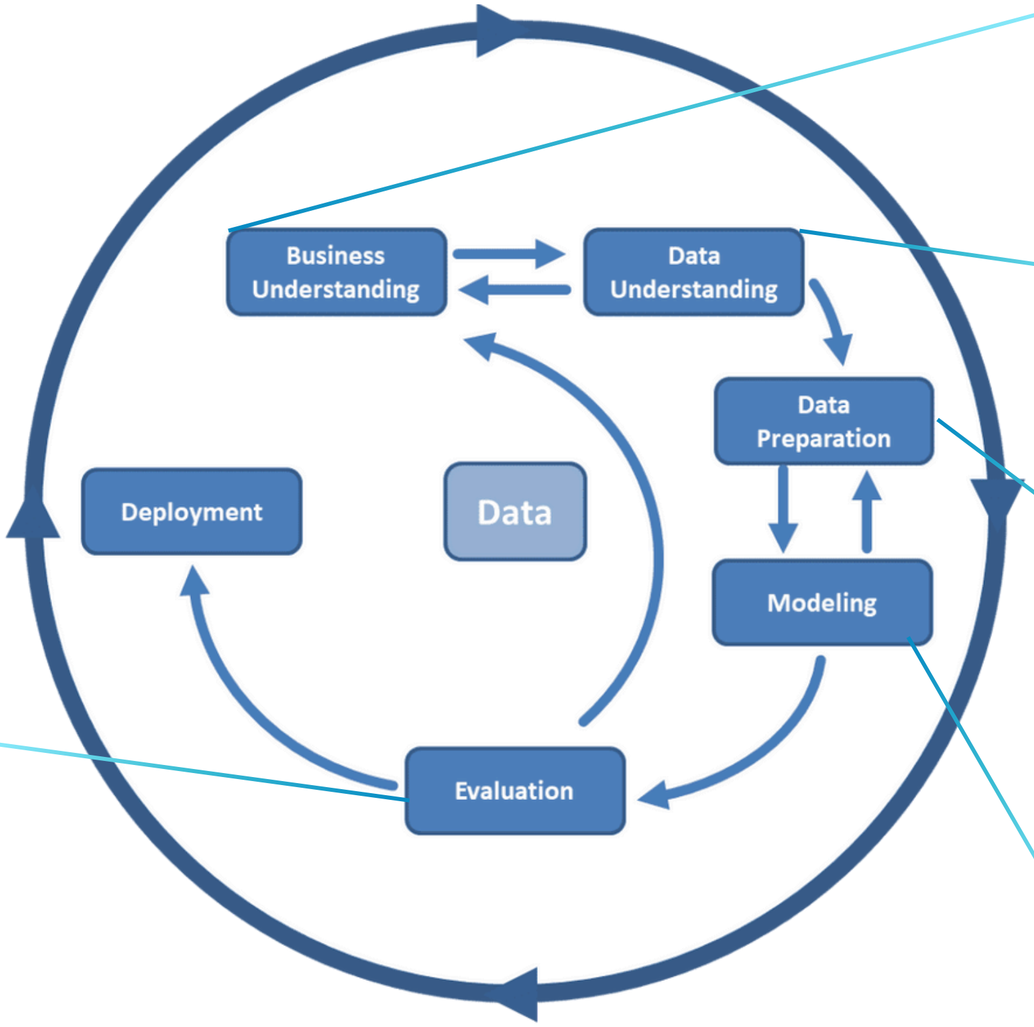
- Data cleaning
- Feature engineering
- Label encoding for performance categories

4. Modeling

- Train & Test multiple models: Logistic Regression, Random Forest, XGBoost, QRM
- Build multiple ML Pipelines

Source: (Muhammad Fawwaz Narendra, 2025)

ML Project Life Cycle



1. Business Understanding
- To predict employee performance class using SMART KPI.

2. Data Understanding
- Explore dataset structure, feature, distributions, and anomalies.

3. Data Preparation
- Data cleaning
- Feature engineering
- Label encoding for performance categories

4. Modeling
- Train & Test multiple models: Logistic Regression, Random Forest, XGBoost, QRM
- Build multiple ML Pipelines

5. Evaluation
- Assess performance using accuracy, precision, recall, F1-score, Hamming loss
- Select best-performing model → XGBoost

Source: (Muhammad Fawwaz Narendra, 2025)

ML Project Life Cycle

1. Business Understanding
- To predict employee performance class using SMART KPI.

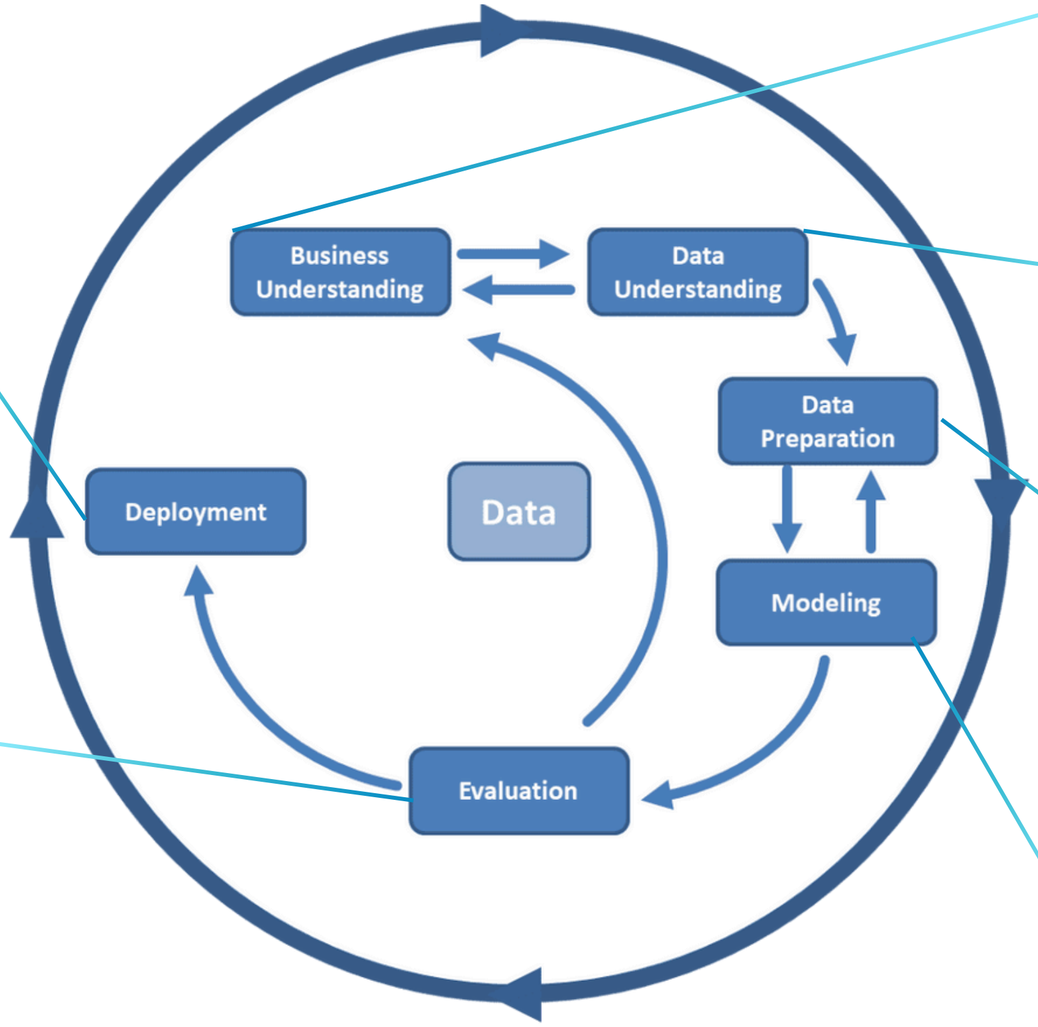
2. Data Understanding
- Explore dataset structure, feature, distributions, and anomalies.

3. Data Preparation
- Data cleaning
- Feature engineering
- Label encoding for performance categories

4. Modeling
- Train & Test multiple models: Logistic Regression, Random Forest, XGBoost, QRM
- Build multiple ML Pipelines

5. Evaluation
- Assess performance using accuracy, precision, recall, F1-score, Hamming loss
- Select best-performing model → XGBoost

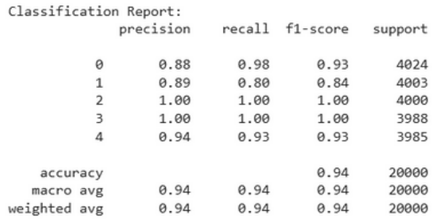
6. Deployment
- Convert final model into a backend service (API)
- Integrate with web interface for file upload → prediction → download → dashboard



Source: (Muhammad Fawwaz Narendra, 2025)

Experiment Plan (ML Pipelines)

Pipeline	Actions	Results

1. Baseline Model	• Trained **initial XGBoost classifier**
	• Used **default** parameters
		`=== Baseline Model Performance ===`
		`Accuracy: 0.9411`
		`Precision: 0.9415`
		`Recall: 0.9411`
		`F1 Score: 0.9403`
		
		Classification Report:
		 

	precision	recall	f1-score	support
0	0.88	0.98	0.93	4024
1	0.89	0.80	0.84	4003
2	1.00	1.00	1.00	4000
3	1.00	1.00	1.00	3988
4	0.94	0.93	0.93	3985
accuracy			0.94	20000
macro avg	0.94	0.94	0.94	20000
weighted avg	0.94	0.94	0.94	20000

Experiment Plan (ML Pipelines)

Pipeline	Actions	Results

1. Baseline Model	• Trained **initial XGBoost classifier**
	• Used **default** parameters
		`=== Baseline Model Performance ===`
		`Accuracy: 0.9411`
		`Precision: 0.9415`
		`Recall: 0.9411`
		`F1 Score: 0.9403`
		
		Classification Report:
		precision recall f1-score support
		0 0.88 0.98 0.93 4024
		1 0.89 0.80 0.84 4003
		2 1.00 1.00 1.00 4000
		3 1.00 1.00 1.00 3988
		4 0.94 0.93 0.93 3985
		accuracy 0.94 20000
		macro avg 0.94 0.94 0.94 20000
		weighted avg 0.94 0.94 0.94 20000

2. Feature Engineering	• Created **derived features**: Salary per hour, Attendance rate
	• Transformed raw KPIs into meaningful **indicators**
		`--- FEATURE ENGINEERING PIPELINE RESULTS`
		`Accuracy: 0.9393`
		`Precision: 0.9396`
		`Recall: 0.9393`
		`F1 Score: 0.9385`
		
		Classification Report:
		precision recall f1-score support
		0 0.88 0.97 0.92 4024
		1 0.89 0.80 0.84 4003
		2 1.00 1.00 1.00 4000
		3 1.00 1.00 1.00 3988
		4 0.94 0.92 0.93 3985
		accuracy 0.94 20000
		macro avg 0.94 0.94 0.94 20000
		weighted avg 0.94 0.94 0.94 20000

Experiment Plan (ML Pipelines)

	Pipeline	Actions	Results

1.	Baseline Model	• Trained **initial XGBoost classifier**
		• Used **default** parameters
			`=== Baseline Model Performance ===`
			`Accuracy: 0.9411`
			`Precision: 0.9415`
			`Recall: 0.9411`
			`F1 Score: 0.9403`
			Classification Report:
			precision recall f1-score support
			0 0.88 0.98 0.93 4024
			1 0.89 0.80 0.84 4003
			2 1.00 1.00 1.00 4000
			3 1.00 1.00 1.00 3988
			4 0.94 0.93 0.93 3985
			accuracy 0.94 20000
			macro avg 0.94 0.94 0.94 20000
			weighted avg 0.94 0.94 0.94 20000

2.	Feature Engineering	• Created **derived features**: Salary per hour, Attendance rate
		• Transformed raw KPIs into meaningful **indicators**
			`--- FEATURE ENGINEERING PIPELINE RESULTS`
			`Accuracy: 0.9393`
			`Precision: 0.9396`
			`Recall: 0.9393`
			`F1 Score: 0.9385`
			Classification Report:
			precision recall f1-score support
			0 0.88 0.97 0.92 4024
			1 0.89 0.80 0.84 4003
			2 1.00 1.00 1.00 4000
			3 1.00 1.00 1.00 3988
			4 0.94 0.92 0.93 3985
			accuracy 0.94 20000
			macro avg 0.94 0.94 0.94 20000
			weighted avg 0.94 0.94 0.94 20000

3.	Feature Selection	• Selected **SMART KPI–aligned** variables
		• Avoided correlation-only filtering due to non-linear relationships
			`--- FEATURE SELECTION PIPELINE RESULTS`
			`Accuracy: 0.9397`
			`Precision: 0.9400`
			`Recall: 0.9397`
			`F1 Score: 0.9390`
			Classification Report:
			precision recall f1-score support
			0 0.88 0.97 0.92 4024
			1 0.89 0.80 0.84 4003
			2 1.00 1.00 1.00 4000
			3 1.00 1.00 1.00 3988
			4 0.94 0.92 0.93 3985
			accuracy 0.94 20000
			macro avg 0.94 0.94 0.94 20000
			weighted avg 0.94 0.94 0.94 20000

Experiment Plan (ML Pipelines)

Pipeline	Actions	Results

1. Baseline Model	• Trained **initial XGBoost classifier**
	• Used **default** parameters
		`=== Baseline Model Performance ===`
		`Accuracy: 0.9411`
		`Precision: 0.9415`
		`Recall: 0.9411`
		`F1 Score: 0.9403`
		Classification Report:
		precision recall f1-score support
		0 0.88 0.98 0.93 4024
		1 0.89 0.80 0.84 4003
		2 1.00 1.00 1.00 4000
		3 1.00 1.00 1.00 3988
		4 0.94 0.93 0.93 3985
		accuracy 0.94 20000
		macro avg 0.94 0.94 0.94 20000
		weighted avg 0.94 0.94 0.94 20000

2. Feature Engineering	• Created **derived features**: Salary per hour, Attendance rate
	• Transformed raw KPIs into meaningful **indicators**
		`--- FEATURE ENGINEERING PIPELINE RESULTS`
		`Accuracy: 0.9393`
		`Precision: 0.9396`
		`Recall: 0.9393`
		`F1 Score: 0.9385`
		Classification Report:
		precision recall f1-score support
		0 0.88 0.97 0.92 4024
		1 0.89 0.80 0.84 4003
		2 1.00 1.00 1.00 4000
		3 1.00 1.00 1.00 3988
		4 0.94 0.92 0.93 3985
		accuracy 0.94 20000
		macro avg 0.94 0.94 0.94 20000
		weighted avg 0.94 0.94 0.94 20000

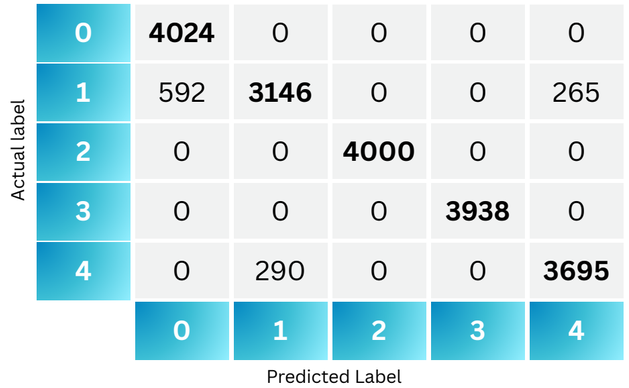
3. Feature Selection	• Selected **SMART KPI–aligned** variables
	• Avoided correlation-only filtering due to non-linear relationships
		`--- FEATURE SELECTION PIPELINE RESULTS`
		`Accuracy: 0.9397`
		`Precision: 0.9400`
		`Recall: 0.9397`
		`F1 Score: 0.9390`
		Classification Report:
		precision recall f1-score support
		0 0.88 0.97 0.92 4024
		1 0.89 0.80 0.84 4003
		2 1.00 1.00 1.00 4000
		3 1.00 1.00 1.00 3988
		4 0.94 0.92 0.93 3985
		accuracy 0.94 20000
		macro avg 0.94 0.94 0.94 20000
		weighted avg 0.94 0.94 0.94 20000

4. Hyperparameter Optimization	• **Stratified cross-validation**
	• **Tuned**: Number of trees, Tree depth, Learning rate, Subsampling rate
		`Best Hyperparameters (manual search): {'n_estimators': 100, 'max_depth': 5, 'learning_rate': 0.1, 'subsample': 0.8}`
		`Best CV Accuracy: 0.9441`
		`--- HYPERPARAMETER OPTIMIZATION RESULTS ---`
		`Accuracy: 0.9426`
		`Precision: 0.9440`
		`Recall: 0.9426`
		`F1 Score: 0.9414`
		Classification Report:
		precision recall f1-score support
		0 0.87 1.00 0.93 4024
		1 0.92 0.79 0.85 4003
		2 1.00 1.00 1.00 4000
		3 1.00 1.00 1.00 3988
		4 0.93 0.93 0.93 3985
		accuracy 0.94 20000
		macro avg 0.94 0.94 0.94 20000
		weighted avg 0.94 0.94 0.94 20000

Why we Choose XGBoost?

1. Confusion Matrix (Classification Accuracy)

- Accuracy: 94.41%
- Very few misclassifications occur between adjacent performance levels
- Severe errors (e.g., Low → High) are minimal or absent



	precision	recall	f1-score	support
0	0.87	1.00	0.93	4024
1	0.92	0.79	0.85	4003
2	1.00	1.00	1.00	4000
3	1.00	1.00	1.00	3988
4	0.93	0.93	0.93	3985
accuracy			0.94	20000
macro avg	0.94	0.94	0.94	20000
weighted avg	0.94	0.94	0.94	20000

Why we Choose XGBoost?

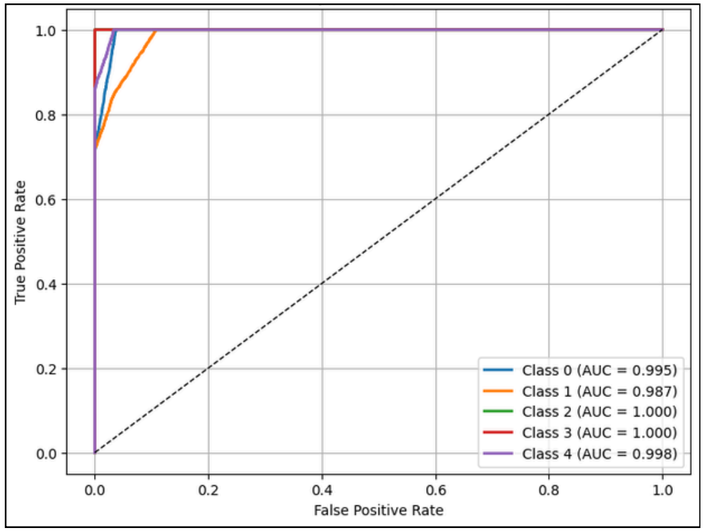
1. Confusion Matrix (Classification Accuracy)

- Accuracy: 94.41%
- Very few misclassifications occur between adjacent performance levels
- Severe errors (e.g., Low → High) are minimal or absent

2. ROC–AUC Performance (Multi-Class Evaluation)

- Low: AUC = 0.9946
- Below Average: AUC = 0.9869
- Average: AUC = 1.0000
- Above Average: AUC = 1.0000
- High: AUC = 0.9976

Actual label	Predicted Label 0	1	2	3	4
0	**4024**	0	0	0	0
1	592	**3146**	0	0	265
2	0	0	**4000**	0	0
3	0	0	0	**3938**	0
4	0	290	0	0	**3695**




```
Classification Report:
              precision    recall  f1-score   support

           0       0.87      1.00      0.93      4024
           1       0.92      0.79      0.85      4003
           2       1.00      1.00      1.00      4000
           3       1.00      1.00      1.00      3988
           4       0.93      0.93      0.93      3985

    accuracy                           0.94     20000
   macro avg       0.94      0.94      0.94     20000
weighted avg       0.94      0.94      0.94     20000
```

Why we Choose XGBoost?

1. Confusion Matrix (Classification Accuracy)

- Accuracy: 94.41%
- Very few misclassifications occur between adjacent performance levels
- Severe errors (e.g., Low → High) are minimal or absent

2. ROC–AUC Performance (Multi-Class Evaluation)

- Low: AUC = 0.9946
- Below Average: AUC = 0.9869
- Average: AUC = 1.0000
- Above Average: AUC = 1.0000
- High: AUC = 0.9976

3. Training vs Validation Loss (Model Stability)

- Training loss and validation loss decrease consistently over boosting rounds
- Both curves overlap closely, indicating:
  - No overfitting
  - No underfitting

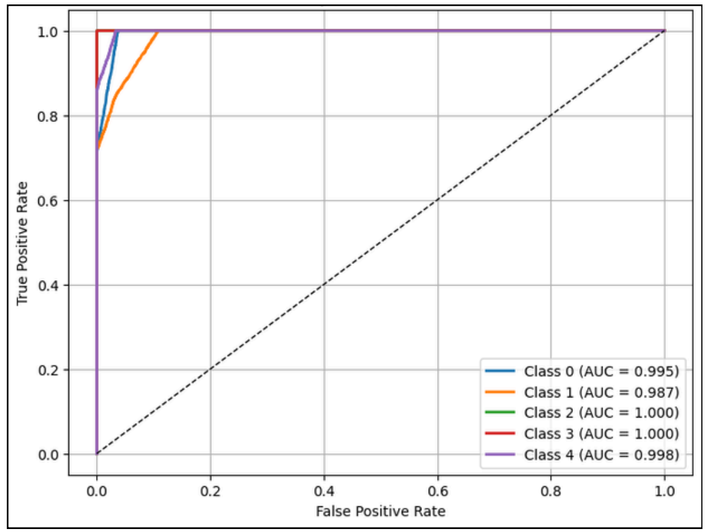
Actual label	0	1	2	3	4
0	**4024**	0	0	0	0
1	592	**3146**	0	0	265
2	0	0	**4000**	0	0
3	0	0	0	**3938**	0
4	0	290	0	0	**3695**

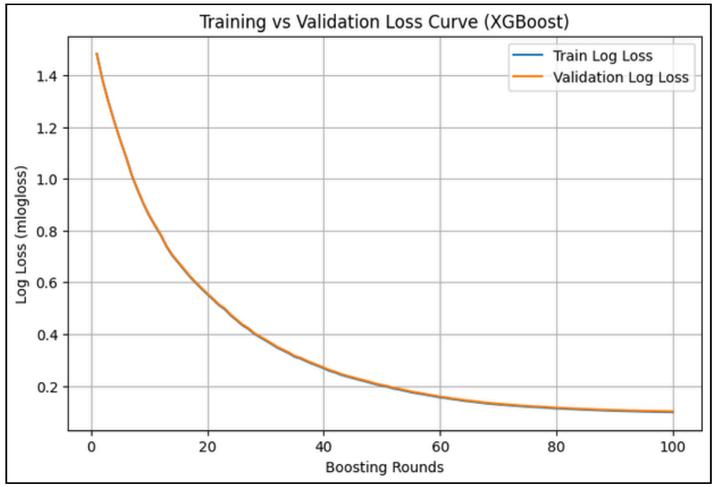
Predicted Label

```
Classification Report:
              precision    recall  f1-score   support

           0       0.87      1.00      0.93      4024
           1       0.92      0.79      0.85      4003
           2       1.00      1.00      1.00      4000
           3       1.00      1.00      1.00      3988
           4       0.93      0.93      0.93      3985

    accuracy                           0.94     20000
   macro avg       0.94      0.94      0.94     20000
weighted avg       0.94      0.94      0.94     20000
```





XGBoost was selected because it achieves **near-perfect ROC–AUC scores**, **strong confusion matrix performance**, and **stable training–validation loss** behavior, ensuring **accurate and reliable performance prediction**.

How XGBoost Works?

Softmax Equation

Probability that employee belongs to class i, given x data

The raw score (logit) that the model produces for class i

$$P(y = i \mid x) = \frac{e^{z_i}}{\sum_{k-1}^{5} e^{z_k}}$$

Normalizes all probabilities so they sum to 1.

Source: (Zhu et al., 2022)

How XGBoost Works?

Softmax Equation

Probability that employee belongs to class i, given x data

The raw score (logit) that the model produces for class i

$$P(y = i \mid x) = \frac{e^{z_i}}{\sum_{k-1}^{5} e^{z_k}}$$

Normalizes all probabilities so they sum to 1.

Example:

Sheldon Cooper

- Projects_Handled: 35
- Overtime_Hours: 33
- Sick_Days: 4
- Training_Hours: 55
- Employee_Satisfaction_Score: 5
- Work_Hours_Per_Week: 48
- Monthly_Salary: RM6300

- Outputs five probabilities, one for each class.
- Class with the highest probability becomes the predicted performance rating.

Class	Probability
Low	0.02
Below Average	0.05
Average	0.05
Above Average	**0.80 ← highest**
High	0.08

Predicted Performance Class = “Above Average”

Source: (Zhu et al., 2022)

Project Impact: ESG

Environmental (E)
Data-driven workforce planning reduces inefficiencies and resource waste

Project Impact: ESG

Environmental (E)
Data-driven workforce planning reduces inefficiencies and resource waste

Social (S)
Objective ML-based performance evaluation reduces bias and improves fairness

Project Impact: ESG

Environmental (E)
Data-driven workforce planning reduces inefficiencies and resource waste

Governance (G)
Transparent metrics and dashboard support accountable decision-making

Social (S)
Objective ML-based performance evaluation reduces bias and improves fairness

8 DECENT WORK AND ECONOMIC GROWTH

9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

Thank You for Your Attention!