

## Model Development Phase Template

Date	30 June 2025
Team ID	NONE
Project Title	Employee Performance Prediction using Machine Learning

### Feature Selection Report Template

The model building phase focuses on developing a regression-based machine learning system to predict employee productivity (actual\_productivity) using historical work data. Three algorithms — Linear Regression, Random Forest, and XGBoost — were trained and evaluated using  $R^2$  score as the primary metric. The Random Forest model demonstrated the best performance ( $R^2 \approx 0.46$ ), effectively capturing non-linear relationships between features like over\_time, idle\_time, and department. The selected model was then saved for integration into a Flask-based web application, enabling real-time productivity predictions

Feature	Description	Selected (Yes/No)	Reasoning
Date	Date of data collection	No	Not relevant for prediction; serves as metadata.
Quarter	Quarter of the year	Yes	Indicates seasonal trends in productivity.
Departement	Department where employee works	Yes	Different departments may have varying productivity levels.
Day	Day of the week	No	Not directly predictive of productivity.

Team	Team identifier	No	Team-specific details not included in analysis.
Targeted_productivity	Targeted productivity goal	Yes	Provides context for actual vs. expected performance.
smv	Standard minute value (efficiency metric)	Yes	Crucial for understanding work efficiency.
wip	Work in progress	Yes	Reflects workload and potential bottlenecks.
Over_time	Hours worked beyond regular hours	Yes	Strongly correlates with productivity
incentive	Bonus or incentive amount	Yes	Motivational factor impacting performance.
Idle_time	Time spent idle during shifts	Yes	Inverse relationship with productivity.
Idle_men	Number of idle employees	Yes	Indicates team inefficiency.
no_of_style_change	Number of style changes	Yes	Team size impacts overall output.
no_of_workers	Total number of workers	Yes	Target variable for prediction.

actual_productivity	Actual productivity achieved	Yes	Target variable for prediction.
---------------------	------------------------------	-----	---------------------------------