Model Development Phase Template

Date	05 August 2025
Skillwallet ID	SWUID20250186419
Project Title	Employee Productivity Prediction Application
Maximum Marks	5 Marks

Feature Selection Report Template

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

Feature	Description	Selected (Yes/No)	Reasoning
date	The date of the day a worker was in the factory.	No	This feature is too granular in its original format. It will be used to engineer new, more meaningful features like day_of_week and month, which are more useful for identifying patterns.
quarter	The quarter of the month.	Yes	This categorical feature can potentially capture seasonal or cyclical productivity patterns. It will be one-hot encoded to be used in the model.
department	The department a worker belongs to (e.g., 'sweing', 'finishing').	Yes	The department is a critical factor that could have a significant impact on productivity. It will be one-hot encoded to be used in the model.
day	The day of the week.	Yes	The day of the week can influence productivity (e.g., lower productivity on Mondays or

			Fridays). This feature will be encoded numerically for the model.
team	The team number of the worker.	Yes	Team dynamics can affect productivity. This feature is important to distinguish between different teams and their performance. It will be treated as a categorical feature and encoded.
targeted_productivity	The target productivity level set by the management for each team.	Yes	This feature is a direct target set by management and will likely be highly correlated with the actual productivity. It's a key predictor.
smv	The Standard Minute Value, which is the time allocated for a specific task.	Yes	SMV is a measure of the complexity of a task and is expected to have a strong relationship with actual productivity. It's a crucial input feature.
wip	Work In Progress, which is the amount of unfinished goods.	Yes	The amount of work in progress can impact the workflow and efficiency of the team. This is a relevant numerical feature.
over_time	The number of hours of overtime put in by a team.	Yes	Overtime can be a factor in productivity, either increasing it due to more work hours or decreasing it due to fatigue. It will be included as a numerical feature.
incentive	The amount of incentive offered to a team.	Yes	Incentives are directly related to worker motivation and are expected to be a strong predictor of productivity. This is a key numerical feature.
idle_time	The amount of time a line was idle.	Yes	Idle time directly indicates a lack of productivity. It's a highly

			relevant feature that should be included.
idle_men	The number of idle workers.	Yes	The number of idle workers is a direct measure of inefficiency and is expected to be a strong negative predictor of productivity.
no_of_style_change	The number of style changes a team had in a day.	Yes	Frequent style changes can disrupt the workflow and decrease productivity. This is a good feature to capture such disruptions.
no_of_workers	The number of workers in a team.	Yes	The size of a team can affect its overall productivity. This is an important numerical feature to consider.
actual_productivity	The actual productivity level achieved by a team.	No	This is the target variable we are trying to predict. It is the dependent variable, not an independent feature for the model.