Project Development Phase
Model Performance Test

Date	30 June 2025
Team ID	LTVIP2025TMID38204
Project Name	**Edututor AI: PERSONALIZED LEARNING WITH GENERATIVE AI And LMS INTEGRATION**
Maximum Marks	10 Marks

THE **Model Performance Testing** template tailored for **Edututor AI: Personalized Learning with Generative AI + LMS Integration**, ready for your team to fill:

---

1. Metrics

Regression Model (e.g., predicting student engagement or test scores):

- MAE –
- MSE –
- RMSE –
- R² score –

Why these metrics matter:

- **MAE** and **MSE/RMSE** quantify average prediction error—the former being more robust to outliers, while RMSE penalizes larger errors heavily arxiv.org+1arxiv.org+1analyticsvidhya.com+15linkedin.com+15researchgate.net+15.
- **R²** shows how much variance in outcomes (e.g., test scores) the model explains linkedin.com+7researchgate.net+7eddys2007.wordpress.com+7.

Classification Model (e.g., predicting mastery level, correct/incorrect):

- Confusion Matrix –
- Accuracy Score –
- Classification Report – (Precision, Recall, F1-score)

These are crucial to understand not only overall success (accuracy) but also balance between false positives and negatives—especially critical in an educational context to avoid mislabeling student abilities .

S.No.	Parameter	Values	Screenshot
1	**Regression Model**	MAE – MSE –	

S.No.	Parameter	Values	Screenshot
		RMSE –	
		R² –	
		Confusion Matrix –	
	Classification Model	Accuracy –	
		Classification Report –	

2. Tune the Model (Hyperparameter Tuning & Validation)

S.No.	Parameter	Values	Screenshot
2	Hyperparameter Tuning	Hyperparameters tuned (e.g., learning rate, depth, regularization) –	
	Validation Method	e.g., k-fold CV (k = 5), stratified CV, hold-out test set –	

Note: Use tuning and validation methods appropriate for Edututor AI's tasks—e.g., stratified sampling for classification of mastery levels, time-aware CV for longitudinal student performance.

✅ Next Steps

- **Fill in each metric** based on your actual model runs.
- **Attach screenshots**: evaluation tables, confusion matrices, classification reports, and tuning logs or validation curves.
- **Add context**: mention dataset split strategies (e.g., train = Jan–Mar 2025, test = Apr 2025).
- **Interpret results** in terms of educational relevance—e.g., "An RMSE of 5 points means predicted quiz scores are off by ~5 marks on average."