Final Year Projects for AI & ML Students PDF

A strong AI and ML final year project starts with a clear problem and a dataset you are allowed to use, compares a simple baseline with a better model, explains errors honestly and ends with a working demo. Good options for 2026 include document Q&A with retrieval, computer vision for safety or agriculture, time series forecasting, speech and language tools for Indian languages, and small AI agents with clear limits.

Key takeaways

- Problem and data come first. A clean dataset and a fair evaluation matter more than a fancy model.
- Each project here lists data, models and tools, main modules, steps, viva points and a rough cost.
- Costs are mainly GPU time, cloud hosting and API usage. Free notebooks such as Google Colab and Kaggle cover most needs. All costs are approximate and vary by city and vendor.
- Always report a baseline, the metric you chose and why, and where the model fails.
- Deploy a simple demo with Streamlit, Gradio or a small web app.

How to choose your AI and ML final year project

1. **Check the data.** Is there a public dataset with a clear licence, or can you collect data ethically? No data means no project.
2. **Check compute.** Can you train on a free GPU notebook, or do you need to fine-tune a small pretrained model instead of training from scratch?
3. **Pick a measurable goal.** Accuracy, F1 score, mean absolute error or latency, against a baseline.
4. **Think about users.** A model inside a usable app is more convincing than a notebook alone.

Beginner-level final year projects

1. Crop disease detection from leaf images

- **Data:** public plant leaf disease datasets (for example PlantVillage on Kaggle), plus a few local photos.
- **Models:** transfer learning with MobileNet or EfficientNet.
- **Modules:** data cleaning, augmentation, training, evaluation, mobile-friendly demo.
- **Viva points:** transfer learning, confusion matrix, why lab images differ from field images.
- **Approximate cost:** Rs 0 on free GPU notebooks (varies by city and vendor).

2. Student performance prediction

- **Data:** public student performance datasets or anonymised college data with permission.
- **Models:** logistic regression, random forest, gradient boosting.
- **Modules:** feature engineering, model comparison, explanation with SHAP, dashboard.
- **Viva points:** class imbalance, fairness, why predictions must not be used to punish students.
- **Approximate cost:** Rs 0 (varies by city and vendor).

3. Fake news or spam detection for short text

- **Data:** public labelled news or SMS spam datasets.
- **Models:** TF-IDF with logistic regression, then a small transformer such as DistilBERT.
- **Viva points:** tokenisation, precision versus recall, dataset bias.
- **Approximate cost:** Rs 0 (varies by city and vendor).

4. House rent price estimator for Indian cities

- **Data:** public rent listing datasets on Kaggle.
- **Models:** linear regression, XGBoost.
- **Modules:** cleaning, outlier handling, model, web form for predictions.
- **Viva points:** MAE versus RMSE, location encoding, data leakage.
- **Approximate cost:** Rs 0 to 500 for hosting (varies by city and vendor).

Intermediate-level final year projects

5. Helmet and number plate detection for traffic

- **Data:** public helmet detection datasets plus your own annotated frames.
- **Models:** YOLO family object detector, OCR for plates.
- **Modules:** annotation, training, detection, plate reading, violation log.
- **Viva points:** mAP, IoU, privacy and legal limits of camera data.
- **Approximate cost:** Rs 0 to 2,000 for extra GPU time (varies by city and vendor).

6. Electricity demand or sales forecasting

- **Data:** public load or retail sales datasets.
- **Models:** seasonal naive baseline, ARIMA, Prophet, LSTM.
- **Viva points:** time-based train-test split, seasonality, why random splits cheat.
- **Approximate cost:** Rs 0 (varies by city and vendor).

7. Speech-to-text and translation for an Indian language

- **Tools:** pretrained speech models (Whisper or AI4Bharat models), translation models.
- **Modules:** audio capture, transcription, translation, word error rate evaluation.
- **Viva points:** WER, accents and noise, model size versus speed.
- **Approximate cost:** Rs 0 to 1,500 (varies by city and vendor).

8. Movie or course recommendation system

- **Data:** MovieLens or a public course rating dataset.
- **Models:** collaborative filtering, matrix factorisation, content-based hybrid.

- **Viva points:** cold start, precision at k, popularity bias.
- **Approximate cost:** Rs 0 (varies by city and vendor).

9. Credit card fraud detection

- **Data:** public anonymised transaction datasets.
- **Models:** isolation forest, gradient boosting with class weights.
- **Viva points:** extreme imbalance, PR-AUC, cost of false negatives.
- **Approximate cost:** Rs 0 (varies by city and vendor).

Advanced-level final year projects

10. RAG-based study assistant for your syllabus

- **Tools:** an LLM API or open model, embeddings, FAISS or Chroma, LangChain or plain Python, Streamlit.
- **Modules:** PDF ingestion, chunking, retrieval, answer with citations, evaluation set of questions.
- **Viva points:** retrieval quality, hallucination checks, chunk size, prompt design.
- **Approximate cost:** Rs 0 to 2,500 for API usage (varies by city and vendor).

11. AI agent that books lab slots or answers helpdesk queries

- **Tools:** an LLM with tool calling, a small database, a web front end.
- **Modules:** intent handling, tool functions (check slot, book, cancel), confirmation step, logs.
- **Viva points:** agent loops, guardrails, why a human confirms actions.
- **Approximate cost:** Rs 500 to 3,000 for API usage (varies by city and vendor).

12. Medical image classification with explainability

- **Data:** public chest X-ray or skin lesion datasets with clear licences.
- **Models:** CNN with transfer learning, Grad-CAM heat maps.
- **Viva points:** sensitivity and specificity, why this is a research demo and not a diagnostic tool.
- **Approximate cost:** Rs 0 to 2,000 (varies by city and vendor).

13. Edge AI on a Raspberry Pi

- **Hardware:** Raspberry Pi with camera.
- **Models:** quantised TensorFlow Lite or ONNX model.
- **Modules:** model compression, on-device inference, latency and accuracy comparison.
- **Viva points:** quantisation, frames per second, power use.
- **Approximate cost:** Rs 6,000 to 10,000 if you buy the board (varies by city and vendor).

14. MLOps pipeline for a deployed model

- **Tools:** MLflow, DVC, Docker, GitHub Actions, FastAPI.
- **Modules:** data versioning, experiment tracking, automated retraining, model serving, drift check.
- **Viva points:** model registry, data drift, reproducibility.
- **Approximate cost:** Rs 0 to 2,000 (varies by city and vendor).

A suggested project timeline

- **Month 1:** problem statement, literature survey, dataset collection and licence check.
- **Month 2:** data cleaning, exploration and a simple baseline model.
- **Month 3:** better models, tuning and error analysis on a held-out test set.
- **Month 4:** demo app, deployment, ethics and limitations section.
- **Month 5:** report writing, plagiarism check, mock viva and buffer for fixes.

Keep a results log with every experiment, its settings and its score. It saves time when you write the report and helps you answer viva questions about why one model beat another.

Project report structure

1. Title page, certificate, declaration, acknowledgement
2. Abstract
3. Introduction: problem, objectives, scope
4. Literature survey of 8 to 15 papers
5. Dataset: source, licence, size, cleaning, splits
6. Methodology: baseline, models, training setup
7. Results: metrics table, confusion matrix or error plots, comparison
8. Deployment and demo
9. Ethics, limitations and future work
10. References

The project report title page generator helps with the first page.

Common mistakes to avoid

- **No baseline.** Without one, 92 percent accuracy means nothing.
- **Data leakage.** Scaling or feature selection on the full dataset before splitting.
- **Wrong metric.** Accuracy on an imbalanced dataset hides poor results.
- **Unlicensed data.** Check the licence and cite the source.
- **Claims beyond the demo.** Say clearly what the model cannot do.

Frequently asked questions

Which is the best final year project for AI and ML students?

One with a clear problem, a licensed dataset, a fair evaluation against a baseline and a working demo. RAG assistants, computer vision for safety or farming, forecasting and Indian language tools are good areas.

Do I need a GPU for an AI final year project?

Not always. Free notebooks like Google Colab and Kaggle give limited GPU time, and fine-tuning small pretrained models needs far less compute than training from scratch.

How much does an AI and ML final year project cost?

Many cost nothing on free tiers. Extra GPU time, API calls or a Raspberry Pi can add a few hundred to a few thousand rupees. These are approximate and vary by city and vendor.

Is using an LLM API acceptable for a final year project?

Usually yes, if your own work is clear: retrieval, evaluation, tools, guardrails and the app. Check with your guide and explain what the API does and what you built.

What do examiners ask in an AI project viva?

Expect questions on your dataset, train-test split, metric choice, baseline, error analysis and the limits of your model.

Is this PDF free?

Yes. Fill in the short form on this page and the PDF downloads right away.

Sources and further reading

- scikit-learn user guide
- PyTorch tutorials
- Hugging Face documentation
- Kaggle datasets
- AI4Bharat

About this PDF

Prepared by the Uncodemy trainer team as a practical project guide. Cloud and API prices change often, so treat costs as approximate. Corrections: info@uncodemy.com.

More free resources and courses: uncodemy.com | info@uncodemy.com | +91 98183 66550