Cleardeals AI Lead Scoring Engine – Intern Screening Assignment

A. Project Snapshot (Scope)

- **Problem:** Brokers spend too much time on low■intent leads.
- **Goal:** Predict high■intent prospects; aim for 2■3× conversion lift.
- **Solution:** AI Lead Scoring Engine delivers an 'Intent Score' inside CRM/WhatsApp in under 300 ms.
- **Key Components:** Data ingestion → pgvector feature store → Gradient Boosted model + LLM re■ranker → FastAPI + Redis → daily retrain & drift watch.
- **Compliance:** DPDP■ready, consent■first, all PII stored in India.

B. 8■Hour Take■Home Assignment

- Show your data thinking, ML basics, and product mindset.
- Work solo. Use bullet points and short sentences.
- Clock starts when you receive this file; submit within 24 hours.

1. Data & Feature Brainstorm – 1.5 h
- List main data source groups (behavioural, demographic, public, third■party).
- For each, suggest 2■3 concrete features and why they help predict intent.

2. Data Quality & Compliance – 1.5 h
- Pick two data■quality challenges and two compliance issues.
- Give one mitigation idea for each.

3. Model Choice & Metrics – 1.5 h
- Explain (in plain words) why Gradient Boosted trees + an LLM re■ranker fit this task.
- Choose two technical and two business metrics to judge success.

4. Improving Lift & Adoption – 1.5 h
- Suggest two ways to boost model lift if performance lags.
- Suggest two tactics to win sales■team adoption if they push back.

5. Real■Time Architecture & Scale – 1 h
- Outline components that keep API latency <300 ms and push scores to CRM/WhatsApp.
- Name one scale risk and how you'd handle it.

6. Monitoring & Continuous Learning – 1 h
- Describe two mechanisms to track drift and keep models fresh.

Submission

- Single PDF; label each section clearly.
- Max 5 pages. Quality > quantity.
- Add your name and LinkedIn/GitHub at top.
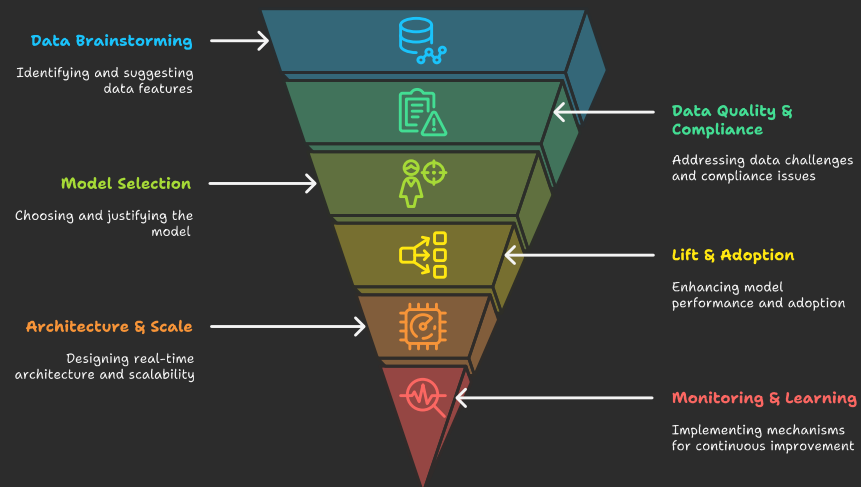- Email back within 24 hours.

How We'll Evaluate You

- Did you cover every task and follow instructions?
- Depth and practicality of ideas.
- Creativity in problem■solving.
- Clear, concise writing.

Expert Tips

- Tie every technical choice to business impact.
- Assume imperfect data – propose workarounds.
- Keep it skimmable – we have 250 applications.
- State assumptions when unsure.

AI Lead Scoring Engine Development Funnel

Data Brainstorming
Identifying and suggesting data features

Data Quality & Compliance
Addressing data challenges and compliance issues

Model Selection
Choosing and justifying the model

Lift & Adoption
Enhancing model performance and adoption

Architecture & Scale
Designing real-time architecture and scalability

Monitoring & Learning
Implementing mechanisms for continuous improvement

Made with Napkin

AI Lead Scoring Engine Development Timeline

1.5 hours
Data & Feature Brainstorming

1.5 hours
Data Quality & Compliance

1.5 hours
Model Choice & Metrics

1.5 hours
Improving Lift & Adoption

1 hour
Real-Time Architecture & Scale

1 hour
Monitoring & Continuous Learning