**🎥 Suggested Video Structure (20 minutes)**

**1. Introduction & Project Context (2 mins)**

* Quick self-introduction and project title: *DeepCSAT – Predicting Customer Satisfaction in E-commerce*.
* Why CSAT is important for e-commerce businesses (loyalty, retention, growth).
* Problem statement: traditional surveys vs predictive, real-time ML models.

**2. Dataset Overview (3 mins)**

* Source: Shopzilla dataset, one-month customer interactions.
* Features: structured (order, timestamps, agent, product, handling time) + unstructured (customer remarks).
* Target: CSAT score (discrete integer).
* Key data challenges: mixed datatypes, missing values, categorical variety, noisy text.

**3. Data Cleaning & Feature Engineering (4 mins)**

* Handling missing values, duplicates, irrelevant fields.
* Derived features: response delay (issue reported → issue responded), item price binning, missing indicators.
* Encoding categorical variables.
* Processing Customer Remarks: text cleaning + embeddings/TF-IDF.
* Emphasis on consistency: preprocessing pipeline saved with joblib for inference.

**4. Model Development (4 mins)**

* Why ANN (captures nonlinear interactions, flexible with mixed features).
* Architecture: input layer (structured + text), hidden layers with ReLU + dropout, output softmax for CSAT classes.
* Training details: optimizer (Adam), learning rate schedule, early stopping, stratified split.
* Handling class imbalance (weights/oversampling).
* Alternatives considered (regression, ordinal models) — why classification chosen.

**5. Evaluation & Insights (3 mins)**

* Metrics used: Accuracy, F1-score, MAE (for ordinal awareness).
* Confusion matrix → where model makes mistakes (usually adjacent scores).
* Key drivers identified with SHAP: response delay, product category, agent tenure, remarks sentiment.
* Business takeaways: faster responses & coaching for specific shifts/agents → higher CSAT.

**6. Deployment & Streamlit App Demo (2 mins)**

* Show the app UI (walk through screenshot or live demo).
* Highlight: user selects interaction details → model predicts CSAT instantly.
* Explain artifact loading (preprocessing pipeline, trained ANN, label encoder).
* Reproducibility: versioning, requirements.txt, GitHub repo.

**7. Limitations, Ethics & Future Work (1 min)**

* Dataset limited to one month, imbalance in some categories.
* Predictions should guide service improvement, not penalize agents.
* Future: transformer embeddings for text, ordinal regression, integration with live CRM systems.

**8. Conclusion & Call-to-Action (1 min)**

* Recap project goal and outcomes.
* Business impact: proactive improvement in customer service.
* Point to GitHub repo & notebook for reproducibility.
* End with thanks + contact details if required.