📘 End-to-End Stock Price Prediction System (TCS & Other MNCs)

🔷 1. Data Pipeline (Backend ML Workflow)

Module 1: Financial Data Acquisition

- Fetch stock data using **Yahoo Finance API**
- Save raw dataset → TCS_raw.csv

Module 2: Data Cleaning & Feature Engineering

1. **Handle Missing Values** → Forward-fill, Backward-fill
2. **Handle Outliers** → Boxplots, IQR clipping
3. **Feature Engineering**
    - SMA (10, 50)
    - EMA (20)
    - RSI (14-day)
4. Save dataset → TCS_cleaned_features.csv

Module 3: Data Preparation

- Convert date → ordinal
- Select features (Close, High, Low, Open, Volume, SMA, EMA, RSI)
- Normalize using **MinMaxScaler**
- Save scalers → scaler_X.pkl, scaler_Y.pkl

Module 4: Models

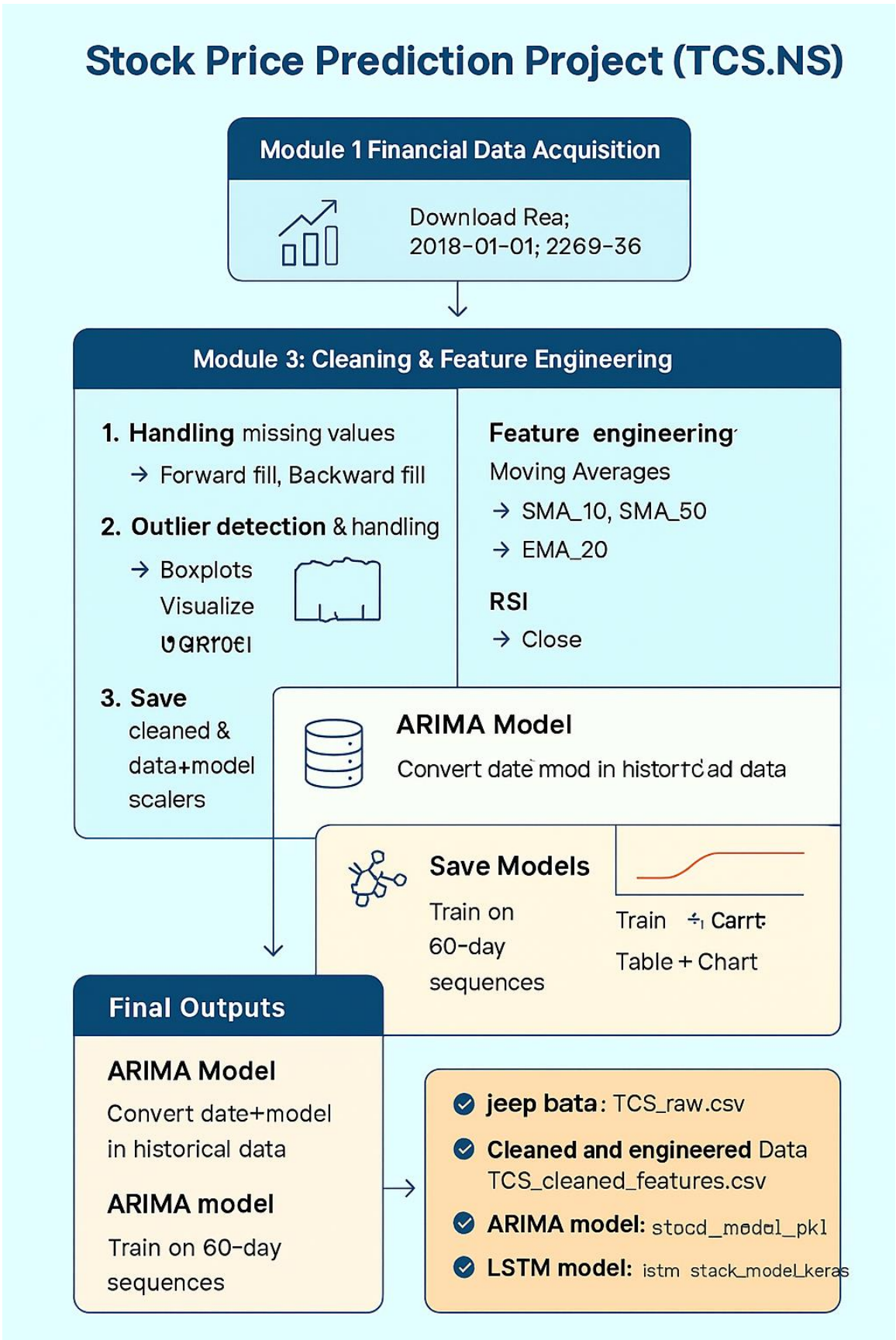
◈ ARIMA Model

- Time-series model on Close price
- Forecast future values till 2027
- Save model → stock_model_arima.pkl

◈ LSTM Model

- Neural network trained on 60-day closing price sequences
- Train-Test split → 80/20
- Trained for 20 epochs
- Save model → lstm_stock_model.keras

Stock Price Prediction Project (TCS.NS)



🔷 2. Streamlit Application (Frontend for Users)

Step 1: User Authentication

- 🔑 Login required
- Users:
    - saurabh / 12345
    - admin / admin123

Step 2: Select Company

- Dropdown list of top Indian MNCs (TCS, Infosys, Reliance, HDFC, SBI, etc.)
- Downloads last **5 years** of stock data from Yahoo Finance

---

Step 3: Prediction Setup

- User selects **future date**
- App calculates prediction steps = (future date – last available date)

Step 4: Choose Mode

1. **Single Model Mode**
    - Select **ARIMA** or **LSTM**
    - Output:
        - Predicted closing price on selected date
        - Last close price
        - 📈 / 📉 Trend + % change
        - Chart: Historical vs Forecast
2. **Comparison Mode**
    - Runs both **ARIMA & LSTM**

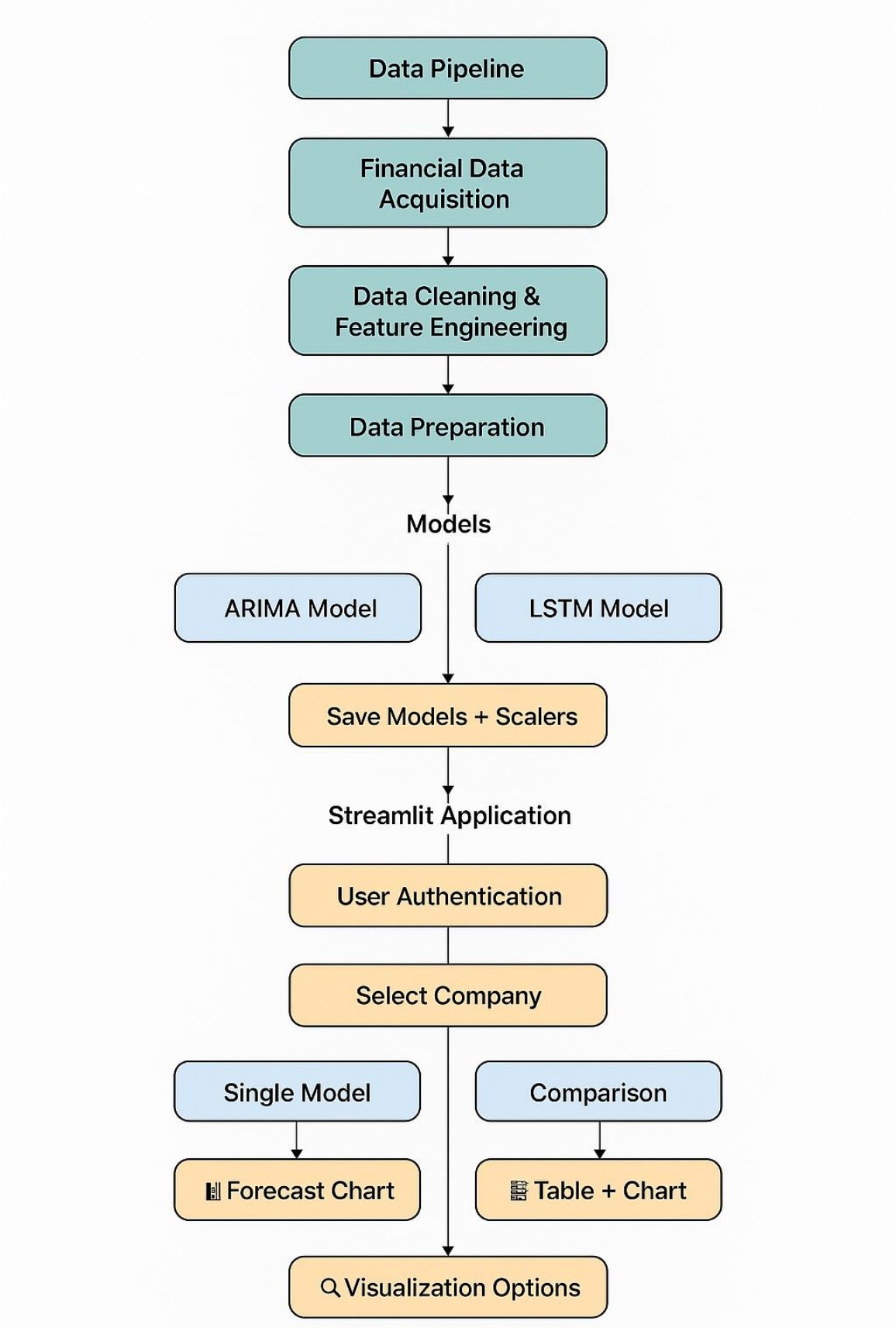
- Outputs:
    - Table with predicted prices & % change
    - Highlights **best model** (closest to last close)
    - Chart: Historical + Both forecasts

Step 5: Visualization Options (Sidebar)

- **None** → Only predictions
- **Moving Averages** → SMA 10 & SMA 50 with closing price
- **Correlation Heatmap** → Feature correlation (Open, High, Low, Close, Volume)

🔷 3. Final Outputs

- 🗂️ Files generated:
    - TCS_raw.csv
    - TCS_cleaned_features.csv
    - stock_model_arima.pkl
    - lstm_stock_model.keras
    - scaler_X.pkl, scaler_Y.pkl
- 📊 App features:
    - Single or Comparison forecast
    - Interactive charts
    - Trend analysis
    - Moving averages & correlation heatmap


Data Pipeline
Financial Data Acquisition
Data Cleaning & Feature Engineering
Data Preparation
Models
ARIMA Model
LSTM Model
Save Models + Scalers
Streamlit Application
User Authentication
Select Company
Single Model
Comparison
Forecast Chart
Table + Chart
Visualization Options