**1. Data Preparation**

* **Objective**: Preprocess and feature engineer 1-year hourly miner flow data.
* **Tools**: CryptoQuant
* **Deliverables**:
  + Cleaned and processed dataset
  + Engineered features for model training
  + Labeling for Buy/Sell/Hold predictions

**2. CNN Model Development**

* **Objective**: Develop a Convolutional Neural Network (CNN) model to predict Buy/Sell/Hold signals.
* **Deliverables**:
  + Trained CNN model
  + Performance evaluation using classification accuracy and confusion matrix

**3. Backtesting the CNN Strategy**

* **Objective**: Simulate trading based on CNN predictions to evaluate strategy effectiveness.
* **Tools**: Custom backtest function
* **Metrics**:
  + Sharpe Ratio: ≥ 1.8
  + Max Drawdown: ≥ -40%
  + Trade Frequency: ≥ 3% of rows
* **Deliverables**:
  + Backtest results with performance metrics
  + Forward test using 2024 data

**4. Hidden Markov Model (HMM) Integration**

* **Objective**: Integrate HMM for market regime detection (bull, bear, sideways).
* **Tools**: hmmlearn package
* **Approach**:
  + Train HMM on key features (flow, volatility, etc.)
  + Label each time step as a market regime
  + Use regime as an additional feature for CNN or build separate models for each regime
* **Deliverables**:
  + Regime-labeled dataset
  + Updated CNN model with HMM integration or separate regime models

**5. Data Expansion with Additional Sources**

* **Objective**: Enhance the model by adding more data sources.
* **Sources**: Glassnode, Coinglass
* **Features to Add**:
  + Glassnode: exchange inflow/outflow, active addresses, SOPR
  + Coinglass: funding rate, long/short ratios, liquidations
* **Deliverables**:
  + Merged dataset with new features
  + Updated feature engineering and scaling process

**6. Optional — NLP Signals Integration**

* **Objective**: Integrate market sentiment from Twitter/news/Reddit.
* **Tools**: LLM APIs (ChatGPT or others like LunarCrush, Santiment)
* **Approach**:
  + Classify market sentiment hourly
  + Output a sentiment score per timestamp
* **Deliverables**:
  + Sentiment score as a feature
  + Updated model with sentiment analysis

**7. Success Metrics Check**

* **Sharpe Ratio**: ≥ 1.8 (Achieved by adding regime detection and tuning risk control)
* **Max Drawdown**: ≥ -40% (Achieved by adding stop-loss and filtering logic)
* **Trade Frequency**: ≥ 3% per row (Achieved by tuning signal smoothing and thresholds)

**8. Final Goal — Wrap it All Together**

* **Objective**: Create a comprehensive pipeline and demonstrate the results.
* **Deliverables**:
  + Complete pipeline from raw data to alpha signals
  + Comparison of CNN-only vs CNN+HMM models
  + Performance metrics and trading charts
  + (Bonus) Clean notebook or Streamlit app for demo