UBS Quantitative Research & Trading - Advanced Chooser Option Pricing Model with Real-World Data & Machine Learning

Department: Quantitative Research & Trading | Duration: 8 Weeks (2 Months)

Project Overview

At UBS, client-driven solutions and proprietary trading strategies often require navigating complex, uncertain market regimes. A **Chooser Option** grants the holder the right to decide, at a future *choice date*, whether the option will be a standard European call or put, with a common strike and final expiration. This instrument is a fundamental building block for:

- **Structured Products:** Creating flexible, cost-efficient portfolios for clients with strong directional views but uncertain about volatility or timing.
- **Risk Management:** Hedging portfolios against asymmetric, regime-dependent risks (e.g., post-earnings announcements, FOMC decisions).
- **Proprietary Strategies:** Serving as a testbed for advanced numerical methods and model calibration, which are core to our QR desk's function.

This project extends the core research of chooser option pricing paper, aiming to **enhance pricing accuracy, adapt to dynamic market conditions, and improve practical applicability** by integrating real-time data collection, feature engineering, and machine learning (ML) optimization. The project bridges academic modeling (Black-Scholes, BSM) with industry practice, empowering interns to build a data-driven, scalable pricing tool for chooser options.

Project Objectives

1. Reproduce and validate the original BSM-based chooser option pricing model using updated real-world data.
2. Collect multi-dimensional real-time/historical data to enrich pricing features (beyond the paper’s basic parameters).
3. Develop ML models to optimize volatility prediction and option pricing (addressing BSM’s "constant volatility" limitation).
4. Compare ML-enhanced models with traditional BSM to quantify performance improvements.
5. Deliver a deployable pricing tool and actionable insights for investors/traders.

Project Timeline & Core Modules

Week	Core Module	Key Tasks	Deliverables
Week 1	**Data Source Design & Initial Collection**	1. **Data Requirement Specification**: Define scope for financial, macroeconomic, and sentiment data 2. **API Setup & Testing**: Configure and test Yahoo Finance, Alpha Vantage, FRED APIs 3. **Initial Data Pull**: Collect JPM stock prices (daily), Treasury rates, VIX for 2018-2024 period	1. Data specification document 2. Working API connection scripts 3. Initial raw dataset (2018-2024)
Week 2	**Data Preprocessing & Feature Engineering**	1. **Data Cleaning**: Handle missing values (interpolation), outliers (IQR), time alignment 2. **Feature Engineering**: - Traditional: Rolling volatilities, daily returns, dividend growth - Advanced: VIX-JPM correlation, interest rate momentum, sentiment scores (0-1) 3. **Pipeline Development**: Build automated preprocessing pipeline with GitHub Actions scheduling	1. Cleaned structured dataset (CSV/Parquet, ≥10 features) 2. Automated preprocessing pipeline code 3. Feature engineering documentation
Week 3	**Original BSM Model Replication**	1. **Model Implementation**: Code the BSM-based chooser option pricing model from the paper 2. **Parameter Configuration**: Set strike=$150, T2=1 year (matching paper parameters) 3. **Initial Validation**: Compare simulated outputs with paper's reported results	1. Fully functional BSM model code (Jupyter Notebook) 2. Model validation notebook with initial comparisons 3. Parameter configuration file
Week 4	**Baseline Model Performance Evaluation**	1. **Error Metric Calculation**: Compute MAE, RMSE between model predictions and actual CME transaction prices 2. **Limitation Analysis**: Identify failure modes in high-volatility periods, sentiment impact gaps 3. **Benchmark Establishment**: Document BSM model performance baseline for comparison	1. Model validation report with error metrics 2. Performance benchmark documentation 3. Optimized, well-commented BSM code
Week 5	**Machine Learning Model Design & Implementation**	1. **Two-Approach Architecture**: - Approach 1: ML volatility prediction (LSTM/RF/XGBoost) + BSM pricing - Approach 2: End-to-end supervised pricing (Linear Regression/GBDT/NN) 2. **Feature Preparation**: Create time-series split datasets (70%/15%/15%) to prevent look-ahead bias 3. **Model Framework Development**: Build initial ML model pipelines	1. ML architecture design document 2. Feature engineering optimization code 3. Initial ML model frameworks (both approaches)
Week 6	**Model Training, Tuning & Comparison**	1. **Hyperparameter Optimization**: Grid/random search with cross-validation 2. **Model Training & Evaluation**: Train final models, evaluate on test set 3. **Performance Comparison**: Compare ML models vs. BSM baseline across MAE, RMSE, R^2 4. **Interpretability Analysis**: Use SHAP/LIME to explain feature importance and model decisions	1. Trained ML models (pickle files) 2. Comparative analysis report with metrics 3. Feature importance visualizations (SHAP plots)
Week 7	**Advanced Analysis & Tool Development**	1. **Extended Sensitivity Analysis**: - SHAP-based impact quantification of new features (sentiment, VIX) - Extreme scenario testing (50% volatility spike, 2% rate hike) 2. **Tool Framework Development**: Begin building Streamlit/FastAPI application structure 3. **Real-time Data Integration**:	1. Comprehensive sensitivity analysis report 2. Pricing tool prototype (basic UI/functionality) 3. Real-time data integration

		Implement auto-update functionality for market data	module
Week 8	**Tool Finalization & Project Delivery**	1. **Tool Feature Completion**: - Implement dual pricing (BSM + best ML model) - Add error margin displays - Complete visualization dashboard (price trends, sensitivity charts, performance metrics) 2. **Final Report Compilation**: Synthesize all findings into 10-15 page report 3. **Demonstration Preparation**: Record tool demo video, create final presentation	1. Fully deployable pricing tool (GitHub repo with README) 2. Final project report (PDF) 3. Tool demo video (5-10 minutes) 4. Final presentation deck

Key Project Milestones

- **Milestone 1 (End of Week 2)**: Complete data pipeline; deliver structured dataset with ≥10 features
- **Milestone 2 (End of Week 4)**: Validate BSM model; establish performance baseline
- **Milestone 3 (End of Week 6)**: Complete ML model training; determine optimal pricing approach
- **Milestone 4 (End of Week 8)**: Deliver final product suite (tool + report + demo)

Required Tools

- Data: Alpha Vantage API (free tier), FRED, Yahoo Finance, CME Group Historical Data, Reuters News API (free alternative: NewsAPI).
- Code: Python (Pandas, Scikit-learn, TensorFlow/PyTorch, Streamlit), GitHub (version control + Actions for auto-data collection).

Reference Papers

Exploration of JPMorgan Chooser Option Pricing (2021)

Rubinstein, M. (1991). *Chooser Options* (original chooser option pricing derivation).

Hull, J. C., & White, A. (1987). *The Pricing of Options on Assets with Stochastic Volatilities* (addressing BSM's constant volatility flaw).

Gu, S., Kelly, B., & Xiu, D. (2020). *Empirical Asset Pricing with Machine Learning* (ML applications in financial asset pricing).

Bakshi, G., Cao, C., & Chen, Z. (1997). *Empirical Performance of Alternative Option Pricing Models* (benchmarking non-constant volatility models).