
Generating Volatility Surfaces of Index Options Using Variational Autoencoders (VAE)

COS 513 – Spring 2025 – Project Proposal

Zhuolin Xiang Haoqian Zhang Siqing Zou
Princeton University
{zx4614,hz0531,sz6982}@princeton.edu

1 Background and Motivations

Volatility modeling is crucial in financial markets, as it underpins option pricing, risk management, and trading strategies. The implied volatility surface, which represents the market's expectations of future volatility across different strike prices and maturities, is fundamental to these applications. However, existing modeling methodologies have limitations. Parametric models like Heston and SABR rely on specific assumptions that may not generalize well across varying market conditions. Moreover, market-implied volatilities often contain missing or noisy data, requiring robust interpolation and extrapolation techniques. Given the advancements in generative models, methods like VAEs and LDMs provide a compelling alternative by learning a structured, low-dimensional representation of volatility surfaces without explicit assumptions about their shape.

Our project seeks to evaluate how VAEs can enhance volatility modeling by capturing latent structures and improving both interpolation and extrapolation capabilities. Additionally, we aim to analyze how these latent structures reflect underlying market conditions, providing deeper insights into volatility dynamics. Furthermore, we will explore the integration of VAEs with diffusion models, specifically Latent Diffusion Models (LDMs), to assess whether this hybrid approach can generate a more accurate and stable volatility surface. Our code is uploaded at <https://github.com/zhanghq-chn/VAE-VolSurface.git>

2 Methodologies

Variational Autoencoder (VAE) Our project will employ a VAE framework(see fig.1a) to model the SPX implied volatility surface. The VAE will be trained with an encoder-decoder structure, where the encoder maps the volatility surface into a latent space distribution, and the decoder reconstructs the surface from sampled latent variables. The training objective will balance reconstruction loss and Kullback-Leibler (KL) divergence to enforce meaningful latent representations.

We follow the structure in [1] to construct the VAEs in two ways: the **grid-based** approach and the **pointwise** approach(see fig.2).

Latent Diffusion Models(LDMs) To further enhance the performance, we will adopt a hybrid approach. LDMs in [3] apply diffusion-based denoising techniques in the latent space rather than directly on high-dimensional volatility data, so we expect LDMs can improve the quality of generated surfaces(i.e., robust extrapolation, reduced) while maintaining computational efficiency. Performance comparisons will be conducted to assess whether LDM-enhanced VAEs provide a superior alternative to standalone VAE and traditional parametric methods.

Evaluation Once trained, the performance of our model will be evaluated through reconstruction error, interpolation accuracy, and generalization ability across different market conditions. The model’s output will be benchmarked against traditional stochastic models like Heston [4] and SABR [2] to assess its viability as an alternative approach.

3 Data

Source Wharton Research Data Services (WRDS) [Detailed description](#)

Statistics [Add statistics here](#)

4 Expected Challenges

Several challenges are anticipated in our project.

Ensuring data quality One of the primary difficulties is ensuring data quality, particularly in handling missing or noisy volatility values while maintaining an arbitrage-free surface. Prices of illiquid options may not reflect the true market expectation of future volatility. Also, restricted by data source availability, our data does not contain trade prices, while only providing best bid, best ask, and last trade date, which may bring difficulties in the accuracy of estimation.

Hyperparameter tuning Hyperparameter tuning is another challenge, as optimizing the trade-off between reconstruction loss and KL divergence is critical for meaningful latent space representation. Also, we need to choose the distribution and dimension of the latent space in VAEs carefully to achieve best performance with lower computation costs.

Generalization across various market conditions As financial markets exhibit regime shifts that may not be fully captured by historical data, our VAEs model may not generalize to new regimes and it can be computationally intensive to fit new models as market changes instantly.

Explainability of Latent Space Understanding how each latent dimension corresponds to market phenomena is nontrivial. Unlike traditional parametric models where each parameter has an explicit financial interpretation (e.g., volatility of volatility in the Heston model), the latent factors in a VAE are learned in an unsupervised manner, making their financial meaning opaque.

Lack of Data

References

- [1] Maxime Bergeron, Nicholas Fung, John Hull, and Zissis Poulos. Variational autoencoders: A hands-off approach to volatility, 2021.
- [2] Thomas Mazzoni. Implied and local volatility dynamics in the sabr model, 2011.
- [3] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models, 2022.
- [4] Slava Solganik. Volatility surface in the heston model, 2011.

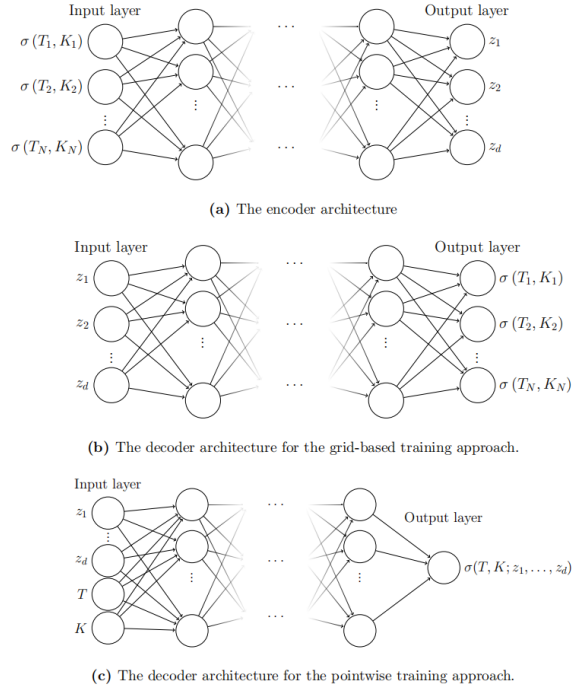
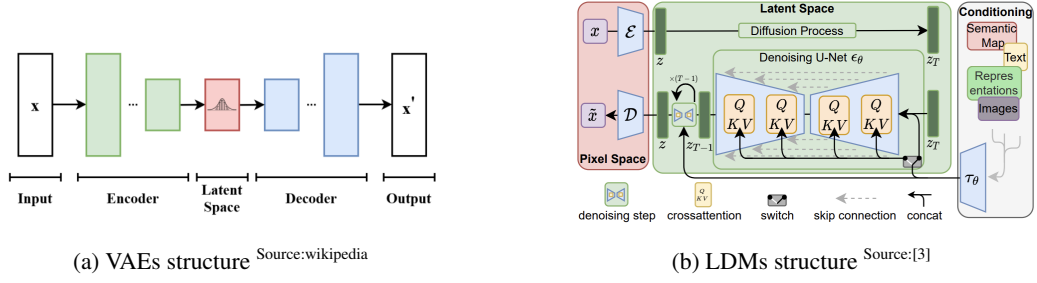


Figure 2: Architectures of grid-based and pointwise method