

Prediction model for Credit Default Swap prices (CDS)

Introduction and statement of purpose

I intend to apply deep learning and large language models (LLMs) to predict the prices of Credit Default Swaps (CDS) in order to develop either hedging, speculating, or arbitrage strategies (depending on the results). Since the 1990s—and particularly after the 2008–2009 financial crisis—the CDS market has attracted growing attention.

In essence, a CDS is a derivate that functions as a form of insurance against the default of a bond¹, whether issued by a corporation or a government. Beyond their role as hedging instruments, CDS contracts are also widely used for speculation and for managing balance sheet risk. Moreover, CDS spreads provide valuable signals about credit conditions:

- Single-name CDS reflect the perceived default risk of individual firms (can be for example Microsoft and for different time bucket .i.e. one, tow ... 30 years).
- Index CDS (e.g., iTraxx Europe, which tracks 125 of the largest European firms) capture broader market sentiment.
- Mortgage-related CDS (e.g., contracts on mortgage-backed securities, MBS) provide insights into the housing market.
- Sovereign CDS measure the creditworthiness of entire countries.

Taken together, CDS markets offer powerful indicators of both firm-level and macroeconomic health.

This study aims to develop models for forecasting future CDS prices. Potential approaches range from standard linear machine learning models and tree-based methods to deep learning architectures. Ideally, a variety of models will be tested in order to identify the most effective

¹ A Credit Default Swap (CDS) is a financial derivative where the buyer pays a regular premium, known as the CDS spread, to the seller. In return, the seller agrees to compensate the buyer if the underlying reference entity (such as a company or sovereign) defaults or experiences another defined credit event. The payout is intended to cover the buyer's losses, typically based on the difference between the bond's face value and its recovery value after default.

framework. Depending on data availability, the empirical focus may be placed on a specific country or on the euro area as a whole.

The feature set will include standard financial and macroeconomic variables such as stock market indices and Treasury yields. In addition, I plan to construct novel indicators by leveraging LLMs to process macroeconomic news (possible also added with other metrics) and transform it into interpretable indices for use in the prediction models. These indices will combine textual and quantitative inputs but will remain grounded in strong economic fundamentals.

Literature review

Even before the creation of the CDO market, Merton (1974) laid the foundation for pricing credit risk. He used company fundamentals to estimate the probability of default, which later informed how insurance against such risks—such as credit default swaps—should be priced.

Avino and Nneji (2014) showed that CDS spreads can be explained or predicted by firm-specific factors, such as assets and liabilities, as well as macroeconomic variables, including the risk-free rate and other yields. Their study employed both linear and nonlinear models to forecast future spreads and demonstrated that these variables can effectively determine CDS pricing. This work provides a foundation for using machine learning and deep learning techniques to identify patterns that may predict future CDS spreads.

More recently, Mao et al. (2023) tested a variety of machine learning and deep learning approaches, with a particular focus on a Merton-LSTM model. They also evaluated alternative models, including standard LSTM networks and Support Vector Machines (SVMs). Their findings indicate that the Merton-LSTM model performs best, achieving the lowest Sharpe ratios and expected losses. This research has paved the way for further development of ML and deep learning methods in the prediction of CDS spreads.

There is already a substantial body of work exploring the use of large language models (LLMs) to enhance the predictive power of economic models or to construct novel economic indices. For example, De Bondt et al. (2025) employed ChatGPT to generate sentiment scores from Purchasing Managers' Index (PMI) commentaries. They found that incorporating these sentiment scores improved GDP prediction models by approximately 20%. Similarly, Silva et al. (2025) leveraged LLMs to convert qualitative central bank statements into quantitative

measures of expected economic developments. Their research demonstrated that backward-looking discussions on topics such as policies and inflation could be used to generate forward-looking indices for factors like inflation expectations and interest rates, or at least to capture central bankers' expectations about these variables. In the financial markets domain, Kirtac et al. (2024) applied LLMs to general economic and financial news to predict stock market movements. Their results were striking, achieving a directional accuracy of 74% and generating impressive year-on-year returns of roughly 100%.

Collectively, these studies suggest that LLMs can be effectively used both to create informative macroeconomic variables and to enhance predictive models. This indicates promising potential for combining machine learning or deep learning approaches to predict market metrics such as CDS spreads while simultaneously leveraging LLM-generated insights to capture current macroeconomic conditions, creating highly powerful hybrid forecasting frameworks.

Data & DMP

No personal data will be used in this project; instead, the focus will be on collecting market data and publicly available economic news, such as central bank statements. The type of CDS analyzed will largely depend on the availability of free online data, which may include single-name CDS, CDS indices, or sovereign CDS. Obtaining suitable text data at daily or weekly frequency may be more challenging. While central banks and other institutions regularly publish statements that should be freely usable, many economic indicators are only released monthly, meaning the scope of the text dataset will depend heavily on both availability and publication frequency. In some cases, web scraping may be required to compile the necessary text data.

Data storage will depend on the size of the dataset. Smaller datasets may be stored locally on my laptop or an external hard drive, while larger datasets will likely require external storage solutions such as CSC's servers.

Methods and Metrics

The modeling for CDS price prediction will involve experimenting with different models and approaches to identify the best-performing one, or at minimum, establish a baseline for what constitutes reasonable model performance. The process will begin with linear machine learning models, followed by tree-based models and SVMs, before progressing to deep learning approaches such as LSTMs. LSTMs appear particularly promising, as the patterns in CDS data are unlikely to be linear given the complexity of the subject.

The target variable will be the CDS price, while the feature set will consist of a wide range of inputs, including risk-free rates, yield curves, stock indices, stock prices, economic variables, credit ratings, the lagged CDS prices, technical indicators such as Larry Williams' %R, and the constructed macroeconomic index. Since the data will be time series in nature, a key challenge will be ensuring sufficient availability of daily or weekly observations.

For the CDS price predictor, success will be evaluated using standard performance metrics such as MSE, R^2 , loss, test accuracy, and other appropriate measures. Multiple models will be tested to identify the best fit, with hyperparameter tuning conducted through methods such as grid search. Data will be split into training, validation, and test sets to ensure robustness of the results.

For the text interpretation task, aimed at constructing an index of economic development, I intend to use a free Ollama model (e.g., LLaMA 3.2) or some embedding model to analyze textual data and integrate the output with macroeconomic variables. The validity of this index will be evaluated based on its consistency with economic theory and support from relevant data.

Both the ML/DL modeling and the text analysis will require significant computational resources. Given the likely limitations of my personal computer, access to external computing resources will be critical. In particular, the use of CSC's Puhti supercomputing environment is expected to play an essential role in enabling these analyses.

Ethics

The primary ethical challenge relates to model transparency, as both deep learning models and LLMs provide limited interpretability. To mitigate this, I will investigate established as well as novel approaches to assess whether the model's results can be meaningfully evaluated and validated. Beyond this, the subject matter itself is largely uncontroversial, and since no human data will be used, no special permits or consent are required.

Timetable

The plan is as follows:

- September – October: Develop a well-defined and focused thesis, finish the research plan, obtain supervisor, gather the necessary data, and determine the models and methods to be applied.
- November – December: Produce initial results in the form of a preliminary model and begin drafting the report.
- January – March: Write the majority of the thesis, aiming to have a near-complete draft by the end of this period.
- April: Focus primarily on refining, polishing, and finalizing the thesis.

Deliverables

The deliverables will consist primarily of a written report presenting the main findings of the thesis. In addition, I will provide all code used for data preparation and collection, the final model for predicting CDS prices, and the constructed macroeconomic index, including its calibration details and the LLM implementation code.

The success criteria will differ between the two main components. For the CDS price predictor, performance will be evaluated using standard metrics, while the macroeconomic index will require a more unconventional assessment. Its validity will need to be demonstrated through consistency with economic theory and alignment with relevant supporting data. The overall performance of the predictive framework will also be evaluated using the Sharpe ratio², if the

² Sharpe ratios are standard methods for evaluating how an investment performed in regards to a risk free rate and the underlying volatility of the investment

pricing models can be leveraged to make trading strategy. The planned pricing/trading strategies will however be more medium term focused (1-3 months).

References

Avino, D., & Nneji, O. (2014). Are CDS spreads predictable? An analysis of linear and non-linear forecasting models. *International Review of Financial Analysis*, 34, 262-274.

de Bondt, G. J., & Sun, Y. (2025). *Enhancing GDP nowcasts with ChatGPT: A novel application of PMI news releases* (No. 3063). ECB Working Paper.

Mao, W., Zhu, H., Wu, H., Lu, Y., & Wang, H. (2023). Forecasting and trading credit default swap indices using a deep learning model integrating Merton and LSTMs. *Expert Systems with Applications*, 213, 119012.

Kirtac, K., & Germano, G. (2024). Sentiment trading with large language models. *Finance Research Letters*, 62, 105227.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2), 449-470.

Silva, T. C., Moriya, K., & Veyrune, R. M. (2025). *From Text to Quantified Insights: A Large-Scale LLM Analysis of Central Bank Communication* (No. 2025/109). International Monetary Fund.