A Deep Learning Approach to Default Risk Migration Matrices with LLM-encoded Data

Yaning You

Department of Statistics & Actuarial Science, University of Western Ontario

Abstract

I will finish this section later, no matter what happens.
Collaborators: should I include anyone's names here?

1 Introduction

This section introduces the scope of the research project. The section will begin with a discussion of the motivation behind the proposed project in terms of application scope and impact. The current practices in industry and its shortcomings will be discussed, which will be followed by an outline of objectives hoped to be achieved through this research. Subsequently the proposal will go into an examination of existing work in predicting default migration matrices and estbalish where the project innovates beyond the current state of the art. Finally, the main data sources to be used throughout the research project will be breifly examined.

1.1 Motivation

Corpoate defaults have implications beyond the company itself. It represents a loss in the company's investors, lenders, and depending on the number of defaults in a time period, the macroeconomy as a whole. Understanding the likelihood of an obligor defaulting helps creditors screen for potential borrowers for more secure returns, investors price instruments at fair value and manage portfolios, and regulators assess the state of the market and decide whether or not to intervene. Thus models of assessing default risk were created by these parties to quantify the risk and track it over time, one of the most popular models being credit ratings.

Credit ratings provide a relatively objective assessment of an obligor's ability to repay its existing debt based off of capital structure, historical performance, internal synergies and external influences on a letter scale. This perception of reliability in turn influences lending capacity, investment instrument pricing, corporate strategy, and regulatory oversight of the obligor as previously mentioned. Credit ratings gained popularity and became one the most common methods of risk assessment due to two factors. Firstly, it is constructed mainly from obligor's financial and economic fundmanetal characteristics, lending the ratings to be highly explicable. More importantly, the ratings are arranged into a letter scale from the highest quality of AAA to companies already in default at D, allowing quick interpretation and decision making, providing a boon for portfolio managers who often face pressure to make high-impact decisions on a tight schedule. Naturally it has been adopted by many in risk modelling and investment management.

However, the Great Financial Crisis of 2008 also exposed several weaknesses of the credit rating system as a investment decision making tool. The letter gradings are still too coarse to assess the credit quality of an obligor outside of screening purposes. Furthermore, credit ratings are assigned relative to other companies within each rating agency; this not only makes aggregation of obligors difficult, but also does not clearly translate to the actual risk of default aside from observing historical rates experienced in each category and performing extrapolations. This means credit ratings are a lagging indicator and are slow to adapt changes in the general state of the economy, particularly changes in the broad environment that may not be directly correlated with the firm, giving the rise in demand for a more granular and objective measure of default risk. Thus Probability of Default (PD) models are born.

1.1.1 Probability of Default models in industry

Unlike credit ratings which assign obligors to a finite, often restrictive, set of credit quality indicators, a PD estimates a concrete probability to the obligor experiencing financial default. This allows rigorous comparisons between firms across agencies as well as scientific examinations of risk management practices. Due to these factors, PD has quickly gained popularity of usage in credit risk modelling, loan pricing, and portfolio management. There are many ways to model PDs, ranging from a simple logistic regression on observed historical defaults to complex multistep models using firm, market and country-level performance over 10+ years. As of the present, all of the "big three" (S&P, Moody's and Fitch) ratings agencies, as well as many financial data service providers (eg. Bloomberg, Morningstar Refinitiv) are offering PD estimates for obligors in its database, a testament to the popularity of PDs in commercial use. The proposal will provide an overview of the PD models of the three agencies and

The PD models for each of the three firms is undisclosed; however there are interesting numances from what information they have revealed. The S&P model uses a dataset of "default flags" (ie. financial statement ratios that can signal default) and derives its PDs with a Maximal Utility approach for optimization and k-fold Greedy Forward for variable selection bounded by AIC for model

parsimony [Alma Chen, 2015]. Fitch employs a Monte Carlo simulation model to estimate the probability and frequency of an obligor's assets falling below its liabilities, mostly relying on historical credit rating data [Ratings, 2013]. Moody's exhibits the most comprehensive and complex model, using macroeconomic indicators on top of individual firm information, as well as incorporating credit migration matrices to calibrate its model over longer periods of time. The PDs appear to be calculated through a censorship-corrected survival analysis model [Moody's, 2018].

Overall, these models use similar data as the credit ratings and employ relatively straightforward methodologies to model the occurrence of defaults within a credit portfolio. This once again lends to highly explicable PD results that might not fully capture the complexities in the market and broad economy. Naturally, more complex methods of modelling PDs emerged. The proposal will introduce the most straightforward, and once again most popular method, the Credit Rating Migration Matrix (CRMM), in the next subsection, and will provide an overview of 1.2.

1.1.2 Credit Rating Migration Matrix

In a portfolio, companies exist in more states than simply default and non-default. The credit rating migration matrix captures a portion of this information by also calculating the probabilities a certain obligor improves or degrades in percieved credit quality. Often the rating migrations are mapped to existing credit rating categories for ease of interpretation. The probabilities of migrating into other categories provide an additional level of granularity into corporate default prediction, providing more stable forecasts into future default occurrences, as well as providing stress testing capabilities on existing portfolios. As such the ability to generate accurate, dynamic CRMMS is crucial to effective pricing and risk management across the financial services industry, motivating the development of highly performant CRMM generating methodologies which will be outlined in 1.2 and 1.3.

1.2 Previous Works

- What has been done? - What gaps in research do you see?

1.3 Objectives

- What do you hope to achieve in this study? - How does this relate to existing research?

1.4 Model introduction

Currently most of the data gathered for this project comes from S&P Capital IQ. As S&P Capital IQ is a commercial platform and access is not widespread across institutions, there Data to train and evaluate the model will be collected in two steps:

1. Construct credit risk data profile for constituents of S&P 500, STOXX 600, and TSX Composite Indices from Capital IQ. The composition of each credit risk data profile is as follows:

   (a) Fundemantal data: this is information regarding individual firms. Data will be collected in these four areas: - Firm volatility (This would involve calculating the % change in the 30 day moving average from previous period) - Equity and Fixed income prices (30 day moving averages) - Financial information (financial statement items and ratios often used to assess borrowing capability) - Total assets - Loss provisions - Effeciency: asset turnover, cash flow cycle - Profitability: Return on equity, return on assets - Leverage: Debt to Equity, EBIT to interest expense - Liquidity: current ratio, tier 1 capital ratio

   (b) Market data: this is information on the underlying index (so the S&P 500, STOXX 600 and TSX Composite). Data will be collected in these areas: - Market volatility (This would be in the form of analyzing the historical performance of the COBOE Volatility Index (VIX) and Euro STOXX 50 Volatility (VSTOXX)) - Index prices

   (c) Macroeconomic data: this is information on the country the index is derived from. Data would be collected in these areas: - real GDP change - Volatility (in the form of interest rates) - Macroeconomic trends (Considering major news events throughout the years)

2. Gather benchmarks to compare the deep learning model output with input. The model would be benchmarked on both accuracy and performance. For accuracy, one source of benchmarks would be the credit migration matrix results presented by the PDiR paper (insert citation here). The model can be trained to also return PDs of individual index constituents, which can be compared with the PDs generated by S&P Capital IQ. Currently the possibility of benchmarking the model performance with S&P generated credit risk migration matrices is under examination; if such data exists and is accessible this will provide another layer of accuracy validation. Comparison of performance will be discussed in 2

2 Methodology

1. Encode the data profiles using LLM into tensors to feed into deep learning model. Current considerations of the transforming LLM models include a Chronos-T5 embedding model to capture temporal relationships, or a General LLM (10B+ parameters) to exploit a more comprehensive knowledge base and can perhaps inference more relevant relationships among the data. Overall, the goal of this step is into transform the firm, equity/debt market and macroeconomic data into a format that can be understood by a deep learning model.

2. Generate predicted credit migration matrices using deep learning model Graph-NN (Horie and Mitsume, 2022) SineNet (Zhang et al., 2024)

3. Compare model results to results produced by existing work If permissible, test efficacy of training one vs. multiple models

References

Giorgio Baldassarri Alma Chen. Pd model fundamentals: Banks - a pioneer model for assessing bank creditworthiness, 2015. URL `https://www.spglobal.com/marketintelligence/en/documents/pd-model-fundamentals-banks-a-pioneer-model-for-assessing-bank-creditworthiness.pdf`.

Moody's. Features of a lifetime pd model: Evidence from public, private, and rated firms, 2018. URL `https://www.moodys.com/web/en/us/insights/credit-risk/features-of-a-lifetime-pd-model.html`.

Fitch Ratings. Fitch portfolio credit model, 2013. URL `https://www.fitchratings.com/fitch-portfolio-credit-model`.