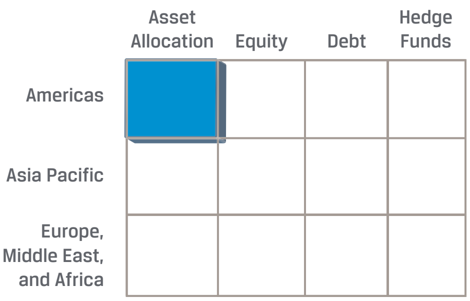
3. REFINING EQUITY TRADING VOLUME PREDICTION WITH DEEP LEARNING: STATE STREET CORPORATION

Contributor: Dajun Wang, CFA[21]



Background

Trading volume prediction is an important topic in finance, especially for those institutions providing asset management or trading execution services. It is important because it not only helps an institution better allocate its trading resources but also gives traders a better view of market conditions. In this trading project, we explore the idea of predicting trading volume with DL technologies and package that process into a service capable of forecasting daily equity trading volumes segmented by markets.

Inputs

A financial institution may serve funds of many investment styles, so the temporal patterns of its daily trading behavior are fairly complex and must be separated by a number of factors, including the following:

1. Financial market indexes. This kind of factor is especially important for active funds. The investment strategy of an actively managed fund usually relies on a set of financial market indexes, such as the CBOE Volatility Index or Merrill Lynch MOVE Index, as inputs. Thus, the movements of these indexes affect the fund's daily trading volume.
2. Market index reconstitution schedule. This kind of factor is likely to affect passive funds. A passively managed fund usually tracks a market index, such as the S&P 500 Index or Hang Seng Index. A reconstitution on these indexes generally results in higher trading volume on the effective day. This phenomenon is widely observed.
3. Historical trading volumes. Historical trading volumes provide the baseline for normal trading levels and the long-term trend.
4. Regional market temporal correlations. Some financial institutions operate around the globe, which includes multiple regional markets. Even though each market operates on its own local holidays and schedules, correlations are still observed among these markets under certain scenarios.
5. Special calendar days. Some calendar days (e.g., Thursday, the last business day in a month, public holidays) have significant influence on daily trading volume.

To provide a predictive service with accuracy that meets a comprehensive financial service provider's requirements, our model has to take all of these factors into consideration. The service ingests these factors daily and returns the predicted trading volume by markets. When the input factors' data accumulate to a certain size, the model retrains itself with the updated data to accommodate potential structural changes.

AI Technology

The core of our predictive service is a sequence-to-sequence model based on a multi-layer convolutional neural network (CNN). The model is able to take in several time series of arbitrary length and output predicted series with the last element as the final prediction.

The model contains two encoder networks and one decoder network. The first encoder applies proper transformation on market index data and extracts features from them. The second encoder does the same to calendar indicators and historical trading volumes. The output of the two encoders is combined and then fed into the decoder network. The decoder network is built with an "attention" mechanism. It contains both CNN layers and dilated CNN layers. The CNN layers in the decoder network are mainly designed to capture short-term information from features, and the dilated CNN layers are designed to capture long-term and periodical information.

The DL framework of Facebook's PyTorch was used to build the model. The implementation is capable of using both CPU and GPU for model training and calculation.

Team Structure and Development Process

Our model development team has two data scientists, one financial engineer, and one business analyst. The team members have education backgrounds that cover mathematics, statistics, machine learning, and quantitative finance.

21 Managing Director, State Street Corporation. This case is prepared based on our interview with him and materials submitted by State Street Corporation.

The project was initiated in August 2018. Since then, the model development team has actively worked together to understand the business background and requirements, discuss potential features and data sources, address data issues, and confirm intermediate results. Review meetings with management occurred monthly. Model developers provided updates and working examples at these meetings to stakeholders, who in turn provided the modeling team with ideas for model enhancements.

When the input factors' data accumulate to a certain size, the model retrains itself with the updated data to accommodate potential structural changes.

In November 2018, a structure of the model was fixed. The modeling team began to wrap the model into a service. In this phase, the model validation team and IT team were involved. The model validation team has one internal model validator who verifies the model's performance under various scenarios. He also advises the model development team on model robustness. The IT team has one architect, two developers, and one quality assurance engineer. They refactor the complete model into a RESTful service in a secure and resilient way. After that, users can access our service with multiple program languages, such as C/C++, Java, Python, Matlab, or with whatever HTTP clients choose, such as internet browsers. Their contributions also include user authentication and authorization, resilient deployment of the model, encrypted data transactions, disaster recovery, and audit logging.

In January 2019, the first beta version of the service was released and tested with State Street internal trading data. The test data set includes daily equity trading volume data covering 63 markets and funds of various trading styles. During the test, real data are fed into our application daily to verify its functionality and prediction accuracy. Our model reached an R2 of 0.86 for the US market and 0.54 for other markets during the test, which is as expected and comparable with the matrix from the training data set. Our service has also successfully foreseen several trade volume spikes triggered by FTSE index reconstitution in the test.

Key Takeaways

Our trading volume prediction project proves that AI technology can be a powerful addition to State Street's front-office core investment functions. It establishes a standard for future projects on converting AI-assisted models into practical and capable tools that drive the bottom line. Future projects will rely on this experience to further improve the company's operation while complying with standards on architecture design, data governance, and project life cycle management.