Teza

White paper – February 2023

All you need is ~~LOVE~~ TRANSFORMERS

Part I

Team Teza in a friendly co-authorship
with ChatGPT ☺

History of Neural Nets in Finance

"A computer will never tell you to buy one stock and sell another… (there is) no substitute … for flair in judgment, and a sense of timing"

Wall Street Journal, 1962

The history of Neural Nets (NN) in finance in many ways follows the general development of *Artificial Intelligence (AI - a term coined by John McCarthy)*. The trading industry first became excited about NN in the 1980s on the coattails of the AI boom in the 1950s-1970s.

Then, AI was all the rage. The theoretical concept laid out by Alan Turing in Computing Machinery and Intelligence was later picked up by Allen Newell, Cliff Shaw, and Herbert Simon who invented the first AI program - Logic Theorist. In 1958 Newell and Simon proclaimed that computers would beat the world chess champion in 10 years. It did happen, of course, but much later – in 1997, when IBM's Deep Blue won a 6-game match against Garry Kasparov.

In 1967, Marvin Lee Minsky prophesied that the problem of AI would be solved in one generation. Years later, in 1975, he further predicted that in "three to eight years we will have a machine with the general intelligence of an average human being." We are still far from that, and it looks like it will take more than one generation from the "G" in AGI (Artificial *General* Intelligence).

In Part I of our Transformer series, we will discuss the three waves of NNs in finance – from early forecasting models to deep NNs – and lay the groundwork for an exciting journey into the world of transformers.

The first wave of Neural Nets - Exploration

While Marvin Lee Minsky's wild claims were not achieved by the 1980s, excitement was high and the money flowed in. Even Wall Street, a bastion of people with "flair in judgment and sense of timing" was not immune to these initial takeover attempts.

During this period, neural networks were mainly used for financial forecasting, such as predicting stock prices or foreign exchange rates. These early models were relatively simple, often using feedforward architectures with one or two hidden layers. They were trained on historical financial data and used to make predictions about future trends. While these models were a significant improvement over traditional statistical methods, they had limitations, such as the inability to handle large and complex financial datasets.

Ultimately, the computing power was too weak, the data was too noisy, and the first systems were too primitive. Along with the rest of the world, traders became disillusioned with Neural Nets.

The second wave of Neural Nets – Computational Optimization

Evolution came in the form of shallow NNs or more deterministic Support Vector Machines (SVM) – a concept by Vladimir Vapnik based on a theory he developed with Alexey Chernovenkis.

The algorithm uses a technique called "kernel trick," which maps the input data into a higher-dimensional feature space, where it is easier to find the maximum margin hyperplane. This enabled the application of SVMs to non-linearly separable data, making it a powerful tool for classification tasks.

In 1992, John Platt proposed a method called Sequential Minimal Optimization (SMO), which greatly improved the efficiency of the SVM algorithm. This made SVMs more practical for large datasets and led to their widespread use in a variety of applications.

Both SMOs and SVMs proved to be some of the more robust statistical methods, and they are still widely used today. Although it was far cry from the promised gizmo that would replace a human trader, it inspired hope about the potential for AI in finance.

The early 2000s marked a significant shift in the development of neural network technology. This period saw the introduction of more advanced neural network architectures and techniques. Neural networks became more sophisticated and were used for a wider range of tasks, such as credit risk assessment and portfolio optimization.

The third wave of Neural Nets – Let's go DEEP!

2012 was a breakthrough year in *deep learning* and *deep neural networks* – starting with the works of Geoffrey Hinton together with Alex Krizhevsky and Ilya Sutskever.

The convolutional neural networks (CNNs) eclipsed the previous state of the art models in image recognition. In the financial industry, CNNs were used for image-based fraud detection and stock price prediction.

Another key advancement was the use of recurrent neural networks (RNNs), which are able to process sequential data, such as time series data. Of those, Long Short-Term Memory (LSTM), introduced in 1997 by Sepp Hochreiter and Jürgen Schmidhuber is still the most widely used. Consequently, RNNs had the ability to handle larger and more complex financial datasets, ultimately leading to more accurate predictions. This development led to a growing number of data companies processing textual and image information, and selling these datasets to algorithmic trading companies.

In addition to these advancements in neural network architectures, the techniques used to train NNs improved as well. One important development was the resurgence of backpropagation algorithms for training NNs. Backpropagation is an efficient way of using the gradient of the loss function to adjust the weights of NN to minimize prediction error. The "backprop" was invented back in the 1960s, but it was believed that by itself it was not enough (persiting until 2010: Breakthrough of supervised deep learning. No unsupervised pre-training. The rest is history). The new ideas – augmentation and dropout made the training more robust.

The trading world was still far away from human replacement, but technology was getting there, and the recovery from the 2008 crisis brought a new wave of capital into AI-driven startups servicing the financial industry.

Enter the transformers

The third wave of neural network usage in finance ushered in a new breed of deep learning techniques – **transformer models**. These models have been trained on massive amounts of data and are able to analyze large and complex financial datasets, such as financial statements, news articles, and social media posts.

A transformer model is a type of neural network that was introduced in a 2017 paper by researchers at Google (Attention Is All You Need). The key innovation of the transformer architecture is the use of self-attention mechanisms, which allow the model to weigh different parts of the input when making predictions. This allows transformer models to effectively process sequences of variable length, such as sentences or paragraphs.

One of the most popular transformer models is BERT (Bidirectional Encoder Representations from Transformers), which has been shown to be highly effective in a wide range of natural language processing tasks, including sentiment analysis, named entity recognition, and question answering. BERT, like many other transformer models, has been trained on massive amounts of text data and has the ability to understand the context in which words are used.

Another transformer model that is now trending in the market – GPT-3 (Generative Pre-trained Transformer 3) – is capable of language translation, summarization, and text generation. GPT-3 is able to generate human-like responses. More importantly, when trained on textual data, it can significantly improve the decision making process of a non-NLP model designed to make trades. GPT-3's applications range from financial services to communications. The power of GPT-3 has been widely recognized around applications such as ChatGPT, which has begun to revolutionize the way many of us communicate and work.

And while college admission offices will struggle to distinguish between ChatGPT's work from a human applicant, there are more practical applications of common sense usage in our field too (LMPriors: Pre-Trained Language Models as Task-Specific Priors).

Even more exciting, transformers, initially used for NLP tasks, happen to be good at mixing data of different modality, such as text, images, and financial data together! (More on this in our next installments).

What's next?

At Teza, we are fascinated with the explosive growth in AI and have a designated team working with Reinforcement Learning (RL), Deep Learning (DL) and Computer Vision (CV). While we do not know when we'll reach singularity, for now, quants and machines happily coexist inside our walls. We believe we can do better with both combined.

We will leave you with that thought and challenge you to guess which portions of this text was written by a machine and which by a human.

Disclosures

This document is provided solely for informational and educational purposes, and there is no consideration given to the specific investment needs, objectives, or tolerances of any recipient. This document is not investment research and should not be treated as such, nor does it represent a formal or official view of Teza. Additionally, Teza's investment positions may, and often will, vary from its conclusions discussed herein based on any number of factors, including client investment guidelines and restrictions. No representation is given that any statements made in this document are accurate or that Teza's objectives will be achieved. This document contains Teza's opinions, and such opinions are subject to change without notice.

This document does not constitute an offer to sell or the solicitation of an offer to purchase any security or investment product (each, a "Product") and should not be relied on in making any investment decision. Any such solicitation or offering may only be made by means of delivery of an approved offering document and relevant subscription documents, all of which must be read in their entirety. No offer to purchase shares in a Product will be made or accepted prior to receipt by the offeree of such documents and the completion of all appropriate documentation. No offer to sell (or solicitation of an offer to buy) will be made in any jurisdiction in which such offer or solicitation would be unlawful.

It should not be assumed that investments described herein will be profitable. Nothing described herein is intended to imply that an investment with Teza is safe, conservative, risk free or risk averse. An investment with Teza entails substantial risks, and a prospective investor should carefully consider the summary of risk factors included in Teza's Form ADV Brochure (and the relevant offering document) in determining whether an investment with Teza is suitable. The risk of loss in trading futures is substantial. This document does not consider the specific investment objective, financial situation or particular needs of any investor and an investment with Teza is not suitable for all investors. Prospective investors should not rely upon this document for tax, accounting or legal advice. Prospective investors should consult their own tax, legal, accounting or other advisors about the issues discussed herein. Historic market trends are not reliable indicators of actual future market behavior or future performance of any particular investment which may differ materially, and should not be relied upon as such. Investors are also reminded that past performance should not be seen as indication of future performance and that they may lose the entirety of their investment. No recommendation is made positive or otherwise regarding individual securities, futures, strategies or other investment products mentioned herein. Information provided about positions, if any, and attributable performance is intended to provide a balanced commentary, with examples of both profitable and loss-making positions; however, this cannot be guaranteed. Certain data and analyses contained herein are based on theoretical and/or backtested model portfolios and are not representative of the performance of accounts that Teza currently manages. The information provided herein is not intended to provide a sufficient basis on which to make an investment decision, and investment decisions should not be based on simulated, hypothetical or illustrative information that have inherent limitations. Unlike an actual performance record, simulated or hypothetical results do not represent actual trading or the actual costs of management and may have under or over compensated for the impact of certain market risk factors. Teza makes no representation that any account will or is likely to achieve returns similar to those shown. Gross performance results do not reflect the deduction of investment advisory fees, which would reduce an investor's actual return. There can be no assurance that any Product advised by Teza will implement the strategies or trading signals referred to herein, or that if implemented any such strategies or signals achieve their investment objectives.

Certain information contained in this document constitutes "forward-looking statements," which can be identified by use of forward-looking terminology such as "may," "will," "should," "expect," "anticipate," "project," "target," "estimate," "intend," "continue," or "believe" or the negatives thereof or other variations thereon or other comparable terminology. Such statements are based on the current expectations and certain assumptions of Teza, and are, therefore, subject to certain risks and uncertainties. A variety of factors, many of which are beyond Teza's control, affect the operations, performance, business strategy and results of the accounts that Teza manages and could cause the actual results, performance or achievements of such accounts to be materially different from any future results, performance or achievements that may be expressed or implied by such forward-looking statements or anticipated on the basis of historical trends.

Tables, charts and commentary contained in this document have been prepared on a best efforts basis by Teza using sources it believes to be reliable, although it does not guarantee the accuracy of the information on account of possible errors or omissions in the constituent data or calculations. Further, the information herein may be superseded by subsequent market events or for other reasons. Teza does not assume any duty to, nor does it undertake to, update the information herein. Charts and graphs provided herein are for illustrative purposes only. No part of this document may be divulged to any other person, distributed, resold and/or reproduced without the prior written permission of Teza.

* * *