Machine Learning in Banking

Fantastic beasts and where to find them

I am an amateur Go player for more than twenty years. The best part of Go is its deep strategic component. The immense board size (19x19) makes the "standard" reasoning and scoring techniques infeasible even on supercomputers – in the year 2008, the best computer Go programs were still mired at just beyond an advanced beginner's level.

In March 2016, I was astonished at the news of AlphaGo winning Go game over 18-time world champion Lee Sedol. Artificial intelligence (AI) relevant headlines kept going these days. Google's image captioning AI claimed to describe photos with 94%+ accuracy. Facebook just tried teaching bots art of negotiation – so the AI learned to lie. Analysts recently forecasted that in 2025, 15% of new passenger car sales worldwide would be autonomous vehicles. No doubt, AI is revolutionizing the world at an unprecedented pace indeed.

More substance than hype

Banking is no exception. AI is impacting banking industry no less than any other. Discussions in the media around the emergence of AI in banking range from the topic of automation and its potential to cut countless jobs to startup acquisitions. Despite the media hype, AI and machine learning (ML) is increasingly embedded into bank's everyday business including chatbots, robo-advisors, intelligent fraud detection tools, robotic trading, and document analysis and classification for legal and compliance.

The depth of knowledge levels required for AI often leads to tremendous efforts and investments for AI projects, and there are limitations and downsides to these AI enabled services as well. Many robo-advisors in banking don't offer common sense financial advice other than rebalancing automation on ETFs. Many of these expert systems did not anticipate the vast amount of implicit knowledge we all share about the world and ourselves – common sense problem is still considered to be among the hardest in all of AI research. Although deep learning – which is driving today's AI explosion – is undeniably mind-blowing, it is still weak in abstraction and reasoning. Humans can learn from very few examples and can do very long-term planning. On the other hand, deep learning requests tremendous data and relatively straightforward pattern recognition. Today's deep learning techniques cannot be scaled to achieve general intelligence – they could only be designed to address a niche problem in banking. Even for a specific task, deep learning appears still "unreliable" and sometimes makes mistakes humans usually don't make (Hackers could purposely attack ML with adversarial examples). And, the transparency of deep neural network used for risk and compliance (such as credit scoring) concerns regulators.

Despite all these limitations, AI is poised for great market expansion and is ready to transform the entire banking sector. AI, one of the three key themes to form the basis for Gartner's report "Top 10 strategic technology trends for 2017", is becoming a critical competitive advantage for banks. The top challenge that keeps bank CIOs up at night right now is about how – not whether – they should invest in artificial intelligence.

Five problems machine learning can solve in banking

In a broad concept, the technology stack of AI includes machine learning, NLP, robots, and many others. A recent McKinsey report[10] says ML and deep learning attracted almost 60 percent of AI investment, and it is an enabler for so many other technologies and applications, such as robotics, NLP and speech recognition. However, for bankers, many confusion of AI could also be traced back to understanding machine learning properly.

There are many ways to classify machine learning relevant technologies – ML is a combination of informatics, analytics, and computer science. Depending on the engineering approach, machine learning tasks are typically classified into a couple of broad categories: supervised learning (labeling data), unsupervised learning (understanding data), reinforcement learning (algorithm learning to react to an environment to perform a certain goal) and others. Machine learning tasks could also be categorized by the outputs from statistics perspective: classification, regression, clustering, dimensionality reduction and others. There are also different classes of algorithms: Linear Regression, Logistic Regression,

AI and ML are ready to become bank's internal combustion engine to power banks' future in coming years

Decision Tree, SVM, CNN, RNN, kNN, GAN, seq-2-seq, and LSTM. All these definitions and classifications – some overlap each other - lead to confusion for banks' senior executives: What are the right use cases? Which are the best-fit AI technologies? Am I deploying an approved mature technology?

Yet CEOs and senior business bankers don't look at machine learning in above ways – it could be a painful experience to explain GAN to a banker. Instead, use cases including the benefits and risks are the common language understood by business users: fraud prevention, risk management, digital assistants, or financial advisors. Unfortunately, there are disconnections between these use cases and the technical terms.

To address that, we recommend to categorize machine learning by what type of real-world problems ML could help to solve in banking: scoring, tactic, labeling, recommending and pricing.

These five problems have successful AI implementation cases in banking and other industries – be it Netflix recommendation or Google photo tags or Uber price hike. Therefore, bank executives are assured not to be guinea pigs. We recommend Bank CIOs to use this classification to explore different niche use cases based on local context – it simplifies bank executives' process to explore AI use cases and saves efforts of consulting from external parties.

Scoring

An objective of scoring is to provide a score - for instance, between 0 and 1 – given a sample with the help of deep learning. One of the basic examples of scoring is the calculated probabilities in a word prediction algorithm that forms the basis for Natural Language Processing (NLP). The same technique has been widely adopted and enhanced by industry, for instance, diagnosing cancer. AI-based credit scoring models which could make sharper predictions of credit risk is a classic "scoring" use case in banking. Loan default rate, customer loyalty, and customer's total relationship value are among other potential use cases.

AI becomes the touch point embedded in every digital channel, the critical factor that will determine its success is how well banks are able to humanize the experience

Pricing

An objective of pricing is to identify the optimal route (price setting) that could map to your customer. Different from traditional rule-based pricing, regression technique can be used to build and forecast a continuous curve of dynamic pricing. Uber uses artificial intelligence to figure out customer's personal price hike. Potential use cases in banking include customer preferential pricing.

Labeling

An objective of labeling is to separate undesirable samples from desirable samples, or vice versa. It uses both supervised (classification) and unsupervised learning (clustering) approach for anomaly detection, for instance, trash email categorization that used by Gmail and Yahoo Mail. Potential classification use cases in banking include transaction fraud detection and document classification. Potential clustering use cases in banking include automated customer segmentation.

Tactic

An objective of tactic is to interact with one counterparty or many of them with a belief to achieve. It may use reinforcement learning to maintain some belief about the counterparty, refine the belief, and then act accordingly. As we all know, game playing is so far the most popular research field to practice tactic techniques. For instance, Google's AlphaGo is now literally the best Go player in the world, and Elon Musk's OpenAI has beaten the world's best Dota 2 players recently. One of the potential "tactic" use cases in banking is trading. For instance, JPMorgan's LOXM has been used in the bank's European equities algorithms business since the first quarter.

Recommending

An objective of recommending is to learn a user's tastes and preferences. Imagine a user x product matrix – some metrics have scored, most entries are missing – and AI could fill in the missing entries. It uses multiple techniques from collaborative filtering, matrix factorization, ensemble learning to latent variable modeling. Researchers started by using supervised learning approach, and are taking reinforcement learning approach recently to build a "feedback system" to continuously enhance the efficiency and accuracy. Netflix and Amazon are among of those famous examples using AI for the recommendation. Potential use cases in banking include financial advice, product recommendation and portfolio recommendation.

The way forward

For banking executives, despite all the challenges, AI and machine learning have become increasingly crucial to make banks keep up with the competition. As every bank has its unique characteristics and context, the best practice for ML investment will differ among banks – it is influenced by different factors including technology, regulation, talents and the bank's local business context. However, based on research and experience, there are four key pieces of advice that will have a major impact on AI implementation decisions.

Despite the media hype, AI and machine learning (ML) is increasingly embedded into bank's everyday business

Identifying niche use cases

Scientists are not yet ready to overcome the limitations of machine learning and proceed toward general artificial intelligence in near future. Instead, there could be an explosion of specific, highly niche artificial intelligence systems in banking.

It looks a platitude to say identifying a right use case is one of the most important things to start with ML – it is true for any new technology, but it is particularly critical for ML. Some of unsuccessful AI implementation is largely due to bringing in an ML platform without knowing the specific problem to solve. However, the breadth and details of ML knowledge are enormous; it has become a hurdle for bank CIOs to evaluate ML use cases effectively. Investment in ML demands more efforts of researching and consulting than it in others - bank executives could use our recommended five problem types that machine learning can solve to conduct a mapping from use cases to the technology stack for a better selection of use cases. An alternative is to seek help from external AI platforms – for instance, Nia from Infosys – with pre-built banking use cases that can fit the bank's context.

Exploring democratized AI tools

An AI enthusiast may lobby IT executives to use Google's Tensor Processing Unit (TPU) replacing Graphics Processing Unit (GPU) to greatly accelerate their neural network computations behind the scenes. But the chip has been specifically designed for Google's TensorFlow framework, and only available to external by Google Cloud (for now). Even if Google makes TPU available to bank's data center in future, there are still concerns: GPU can be scaled for VR streaming and other high-performance computing tasks easily but TPU cannot.

So, many factors need to be considered from selecting AI frameworks to deciding supporting hardware. Fortunately, open source and cloud computing – both are key catalysts for ML and deep learning infrastructure - are now allowing banks to simplify the implementation of deep learning without having to set up or maintain any other infrastructure. For instance, Google's TensorFlow is the most populate framework on GitHub and highly flexible and portable (but maybe slower than others; well, Google doesn't agree on that); Facebook's Pytorch may be faster (but needs better documentation); CNTK makes Microsoft's open source efforts one step further but it is not licensed for commercial use; banks choose Deeplearning4j largely due to the benefits of leverage existing JVM stack and resources (but Java is unpopular in ML research); and, many of these open source platforms are not commercially supported. If banks are more comfortable with commercial support, business platforms like Infosys Nia is an alternative to help bank build an integrated and licensed AI and ML platform.

AI as an experience

Conversational User Interface (CUI) - for instance, chatbot - has become a different type of interface to get banking business done. Best-of-breed vertical chatbots support contextualized conversational experience, by remembering the dialog context and understanding the small talk.

Artificial intelligence will be the main way that banks interact with their customers within the next couple of years. Once AI becomes the touch point embedded in every digital channel, the critical factor that will determine its success is how well banks are able to humanize the experience - friendly chat messages, emojis, happiness, empathy, customer's personal financial plans, and everything else that comes with it - and yet it is as much a design problem as it is a technical problem. In fact, moving forward, this "human touch" experience will be one of the most critical elements that separate a good CUI from a great one.

Addressing regulatory concerns

To use ML specifically in credit scores and models, transparency could be an issue to regulators. Regulators are worried that AI-enhanced credit scores could become a "black box", with full underwriting process not transparent to consumers. Furthermore, machine learning and deep neural networks could make it harder to provide the needed "reason code" to borrowers for credit denial. Bank executives need to work on techniques that would make AI credit-based score decisions more explainable and auditor friendly.

Bottom line

Early this year, we have said AI is going to become the competitive advantage for banks in the future11. Today, despite the media hype that industries are being replaced by machines – AI is indeed a fantastic beast revolutionizing banking at an unprecedented pace. AI and machine learning have demonstrated its ability and potential to redefine customer experience, automate processes more smartly, and manage risks more effectively. AI and ML are ready to become bank's internal combustion engine to power banks' future in coming years. But are banks ready for the changes?

Ethan Wang
Product Manager, Infosys Finacle