



White Paper

The Solution for Difficult Machine Learning Problems: Man and Machine, Together, Continuously Learning

Tazi.ai

353 Sacramento St STE 1800

San Francisco, CA 94111

+1(408)290-5491

info@tazi.ai



Some machine learning problems are more difficult than the others. Even if you have the most powerful machines and a billion instances to learn from, you might not be able to solve them. It is because the difficulty of the problems lies in the fact that the problems are born out of complicated and changing processes with some unobserved variables. In this note, we will present a solution framework for such cases.

Let's take the auto insurance sector as an example. For an auto insurance company, a key metric is the loss ratio (incurred losses divided by collected premiums). It depends on a number of directly observable factors such as the vehicle and driver type and their history, the local traffic, and even the general state of the economy. Insurance policies are then priced by the experts based on the predicted loss ratio for a certain future time period. The company employs insurance pricing experts who keep track of the variations in a limited number of key variables that affect the loss ratio. When they observe changes in these variables, they reflect those changes in the policy prices for certain micro-segments so that the insurance company realizes as little loss as possible.

When predicting the loss ratio via machine learning, the variables that affect the future loss ratio are used as inputs. Particular effect of some of these variables on the loss ratio might be easy to predict. For example due to traffic patterns on a certain intersection, one can predict there might be more accidents there. [Should the insurance company report these findings to the transportation authorities and its customers? This is the subject for another article on machine learning and ethical responsibility.] On the other hand, the effect of the economy is very difficult to predict. Auto insurance policies are priced for a 12 month period and it is very difficult to predict an economic indicator such as the stock market or currency prices so far in the future, even if you use deep neural networks [Gunduz2017]. However, the human insurance pricing expert might have an idea on the fundamental factors that shape the economic conditions for his/her customer segments and make pricing decisions based on those. In short, valuable human expertise exists that can assist in future predictions.



The challenge facing the pricing expert is his/her inability to process millions of policies in terms of hundreds of variables that can take tens of different values. The machine learning models are able to do just that: process massive amount of data in many dimensional spaces and discover patterns in them.

Machine learning models “learn” from past data. However, life changes continuously. The past historical data that trained our machines may not be valid anymore. That is, the many variables and conditions that are embedded in the historical data may have changed. Therefore a machine learning model that learned the patterns in the past data may not be valid for the future. What we need then, are not machine learning models that only learn from historical batch data but rather models that can continuously learn from incoming data.

Of course you could do learning in mini-batches. However, there is no guarantee that the model you learned in this batch will have anything to do with the model you learned on the next batch [REFWhyMiniBatchesAreNotGoodForBusiness]. Machine Learning models are used to take business decisions and actions and business is not only interested in what the model says, but why the model came to that certain conclusion [REFExplainableMLForAutoInsurance]. The zig-zags of your models will negatively affect business continuity and hence decrease the trust and willingness of business to employ machine learning.

Having a continuously learning machine learning system that can learn from many instances and features is necessary but is not enough. It is, in the end, the business expert that makes use of the model results. But the insurance pricing expert carries in his/her mind a lot of experience that could help the models get better. And incorporating this human expertise in the machine learning models requires at least two things: understanding the model results and being able to update them.

The business expert has to understand how machine learning models make decisions. This is where the concept of “explainable AI” or “understandable machine learning” come into play. A key issue here is to tailor the level of explanation to the needs of each person’s level of expertise and daily business activities. There is a lot of work on the right teaching material and activities in education (e.g., see [Poon2010]). For instance the explanation for a business pricing expert can contain: which features are used to decide, what are their importance, what is the reason (pattern) for this decision, how did the model



decide for instances that differ a little bit, is this a loss pattern that existed for a long time or is it a new pattern, what are the policies that show this pattern, what are the features like for this pattern, how many instances have this pattern, etc.

In order to be able to update the machine learning models, the business expert should not be required to have deep machine learning knowledge, nor should she/he need to know the details of the models. She/he can update the models, for example, by overriding a decision made by the machine learning for a certain pattern, changing the confidence for a decision on a pattern, re-labeling noisy instances that exhibit a certain pattern, changing the decision thresholds on a continuous variable, etc.

In addition to all of these, the pricing expert needs to be able to determine the appropriate business actions right where she/he interacts with the model. Those actions can be, for example, reduce the price, have the sales or marketing take a look at these customers, call the agency, and inform the fraud department. The business actions can be across the enterprise and are taken with more efficiency if the other departments also know about how decisions are taken by the machine learning models.

The system described above requires a new paradigm of machine learning understanding and interactions. A system not tailored for the data scientist or machine learning expert, but especially for the business expert. Also, a continuously learning robust system where the models keep learning and producing outputs, no matter how the world changes. And all of these are already here.



ABOUT TAZI

Artificial intelligence (AI) is a source of both huge excitement and apprehension, transforming enterprise operations today. It is more intelligent as it unlocks new sources of value creation and becomes a critical driver of competitive advantage by helping companies achieve new levels of performance at greater scale, growth, and speed than ever before, making it the biggest commercial opportunity in today's fast changing economy.

TAZI is a leading global Automated Machine Learning product/solutions provider with offices in San Francisco. TAZI is a Gartner Cool Vendor in Core AI Technologies (May 2019) and is considered as "[The Next Generation of Automated Machine Learning](#)" by Data Science Central.

WHO WE ARE

Founded in 2015, TAZI has a single mission which is to help businesses to directly benefit from Automated Machine Learning by using TAZI as a superpower, shaping the future of their organizations while realizing direct benefits like cost reduction, increasing efficiency, enhanced (dynamic) business insight, new business (uncovered), and business automation.

WHAT WE OFFER

Through its understandable continuous machine learning from data and humans, TAZI is supporting companies in banking, insurance, retail, and telco industries in making smarter, more intelligent business decisions.



TAZI solutions are based on a most compelling architecture that combines the experiences of 23 patents granted in AI and real-time systems, proven at different global implementations.

Some unique differentiators of TAZI products are:

- Business users can automatically configure custom ML models based on their KPI and the available data. TAZI's Profiler accelerates this process through data understanding and automated cleaning, feature transformation, engineering, and selection capabilities.
- TAZI models learn continuously, and are suitable for today's dynamic, realtime data environments.
- TAZI models are GDPR compliant (no black-box models). They provide an
- explanation in the business domain's terminology for every result they produce.
- TAZI supports multiple (heterogeneous) data sources, i.e.,...: external, batch, streaming, and others.
- TAZI can learn both from human domain experts and from data, which speeds up accuracy improvement.
- TAZI's hyper parameter optimization feature reduces human time spent for model configuration. TAZI products contain algorithms that are developed and coded to be lean, efficient, and scalable.

Tazi.ai

353 Sacramento St STE 1800
San Francisco, CA 94111
+1(408)290-5491
info@tazi.ai