



White Paper

Why Should You Choose Continuous Machine Learning over Batch: Double the Efficiency of Call Center for Bank Marketing

Tazi.ai

353 Sacramento St STE 1800

San Francisco, CA 94111

+1(408)290-5491

info@tazi.ai



SUMMARY

Continuous machine learning technology allows updating the machine learning system at every instance. In a continuous frame of operation, we accept that change is inevitable, models updated automatically at every instance are always ready, make less mistakes and allow more business benefits.

On UCI Bank Marketing data, we compare batch and continuous machine learning and demonstrate the benefit of continuous learning. The continuous machine learning model identifies which customers would accept the offer when called twice more than the batch machine learning model. Using understandable model explanations, we also explain what changes in terms features used and the machine learning model as the data change.

INTRODUCTION

As humans, we keep learning during our lifetime, from everything we are exposed to and at the time we are exposed to. On the other hand, when traditional batch machine learning (batch ML) methods are used, a labeled dataset is collected, model(s) are trained on that dataset. The available data is usually divided into a training and test set. A model is trained on the training set, i.e. the parameters of the model are optimized to fit the training data. Then the trained model is tested on the test set. If the test performance is good, then it is assumed that the performance of the deployed model would also be good. (see i.i.d., independent and identically distributed, assumption). The deployed model predicts the target for every input (instance) it is presented with.

However, as customers' behavior, economy, products, competitors change so does the data. In batch ML since the deployed model is not updated, it goes stale when data changes. Business realizes that the model has more errors than usual, warns the data science team, who try to understand what is happening. Usually there is a wait until enough data is collected. A new model is trained and tested from scratch and deployed. Between the beginning of change and deployment of the new model, errors increase causing revenue losses (See Figure 1). Training and deployment of new models cause pressure on the IT and data science teams, especially if a large number of machine learning models all need to be

Tazi.ai

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

updated at the same time. In some cases, the model rebuilding is a periodic task to keep the models up-to-date. However, re-training of the models can be computationally and time-wise quite an expensive task since the large data set has to be processed each time.

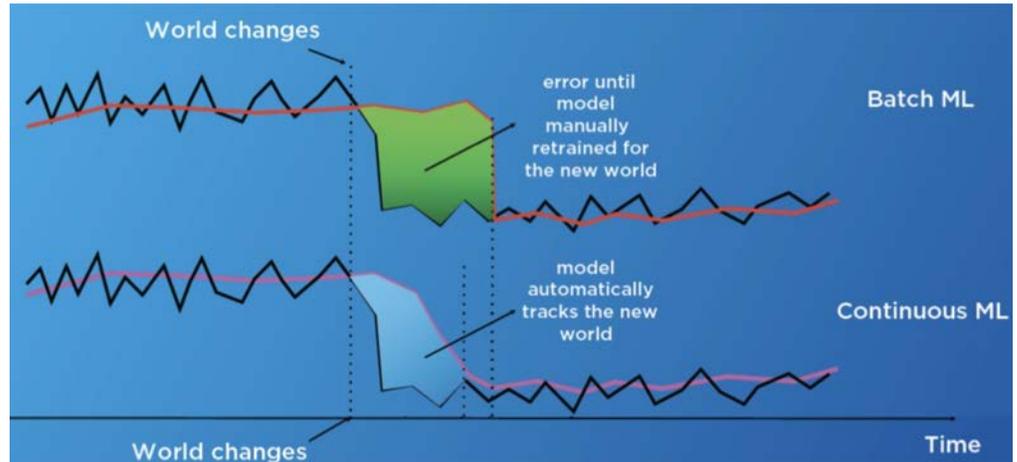


Figure 1: Batch vs Continuous Machine Learning behavior as the world changes.

With continuous machine learning technology, models are updated as soon as data become available. Models are not trained from scratch, but updated incrementally. The update happens automatically and for each instance, without need for data scientist intervention. When the world changes, models learn the new world and adapt to change. Since the machine learning model is updated continuously, the continuous learning model makes less errors than a batch learning model.

Comparison of batch learning and continuous machine learning

We provide comparison of batch and continuous machine learning for the Bank Marketing dataset from the UCI Repository. The dataset contains direct phone marketing campaign results, whether the customer accepted the offer (“yes”) or not (“no”). The dataset has been collected over about 2.5 years and has a clear concept drift [Zliobaite 2016, Gama 2004, Barros 2018]. At the beginning, the dataset contains only around 5% “yes” label, while the “yes” ratio becomes as much as 50% at the end of the campaign (Figure 2). The call center activities became more successful, which meant a change in the ratios of the “yes”, “no”



classes in the dataset. This change in the class ratio is quite significant. At the beginning of the data, we have a class imbalance in the dataset ("yes" is rarely seen). In the end, the classes are quite balanced.

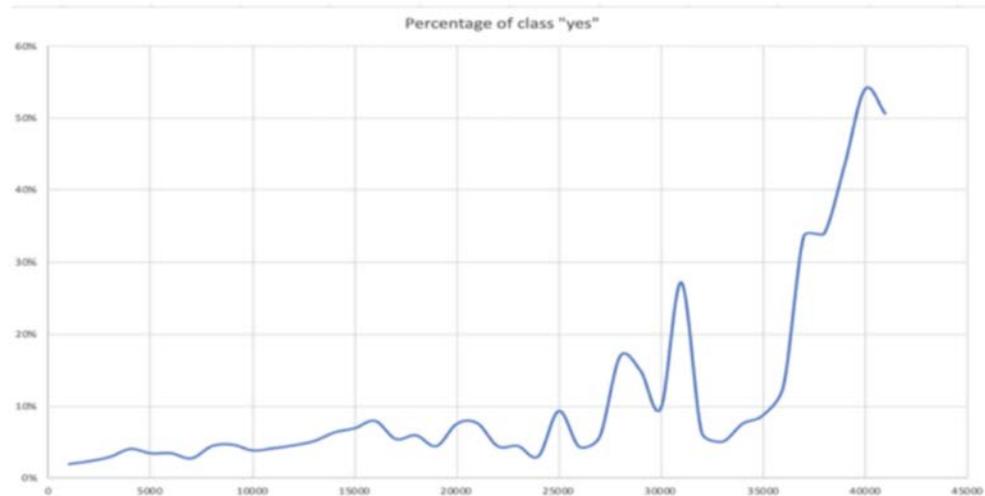


Figure 2: Bank Marketing Data, percentage of class "yes" changes in time.

In order to compare batch and continuous learning, we performed two experiments. In the batch experiment, we trained a model until 30000 instances and stopped updating the model with the new instances and just used the model to create predictions. In the continuous experiment, the model kept training until the end of the data. The model predicted the label for each instance and then was updated with the actual label, i.e. it learned from every instance it predicted. The prediction was made before the model was trained with that instance, therefore the reported results are still test results for the continuous experiment. In a real deployment scenario, think of it as the call results being fed back into the model as they happen.

Since the class ratios change so much, the comparison metric needs to be carefully selected. When there is class imbalance, accuracy is not a good metric. At the beginning of the dataset, we could just predict "no" for all instances and have 95% accuracy. The bank wants to sell the campaign to as many people as possible by calling as small number of people as possible. ML models are expected to predict who should be called by producing a score for class "yes" for all test instances. Then the calls are placed starting with the customer with the highest "yes" score. The amount of actual "yes"s at the highest scored customers shows the performance of the model. We consider the highest, 1, 5, 10, 25%



scored customers and report the percentage of “yes”s found by the model among all the “yes”s in the test set. For comparison, a random model would find only about 1% of “yes”s at the top 1%. Table 1. (and Figure 3) shows the percentage actual “yes”s found by continuous and batch machine learning on the test set. The last column compares continuous machine learning to batch machine learning. Continuous machine learning finds twice or more “yes” customers, customers who purchase when called, than the batch model. This amounts to doubling the efficiency of the call center!

1	%4	%2	%222
5	%20	%8	%256
10	%38	%15	%255
25	%57	%33	%177
50	%71	%58	%122

Figure 3: Bank Marketing Data, comparing top “yes” scored customers’ recall success for Continuous and Batch Machine Learning Models.

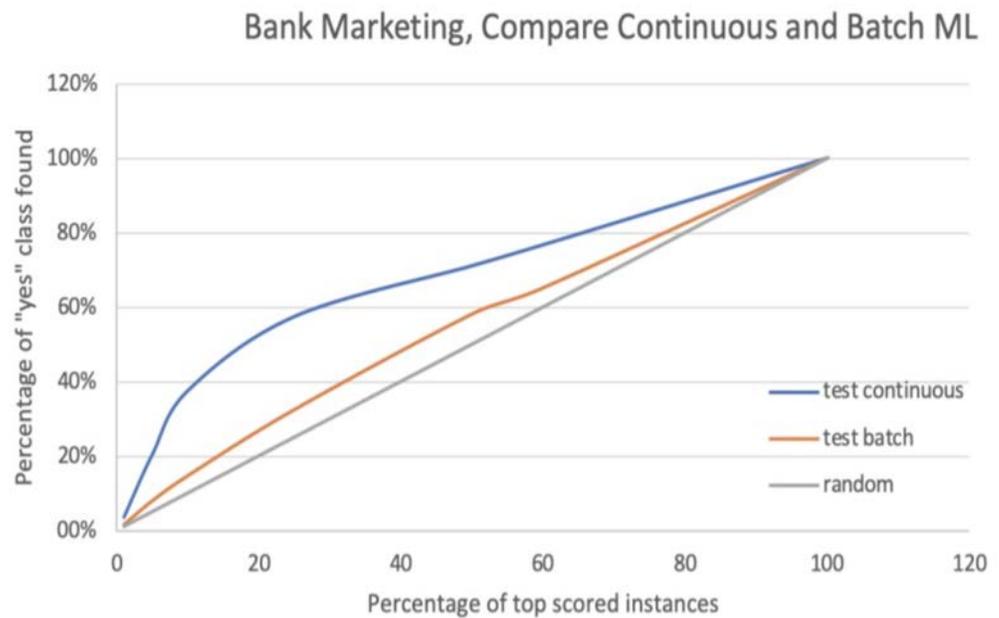


Figure 4: Bank Marketing Data, Continuous Machine Learning discovers customers who will buy more accurately than Batch Machine Learning.

As the final step in comparison of Batch and Continuous Machine Learning, we show the model explanations. Our model explanation (Figure 4.) is a dynamically updated decision tree that explains the underlying ensemble of machine learning algorithms. The tree is interactive, so the call center manager or a data scientist can further inspect it. We also show feature importances. First of all, feature importances between batch and continuous machine learning systems are not too different. Just the order of some features have changed. On the other hand, the explanation models are quite different. Red color shows the customer segments predicted to be in class “no” and the blue shows the “yes”. The continuous machine learning model has learned the customer behavior at the later part of the data and classifies more segments as “yes”.

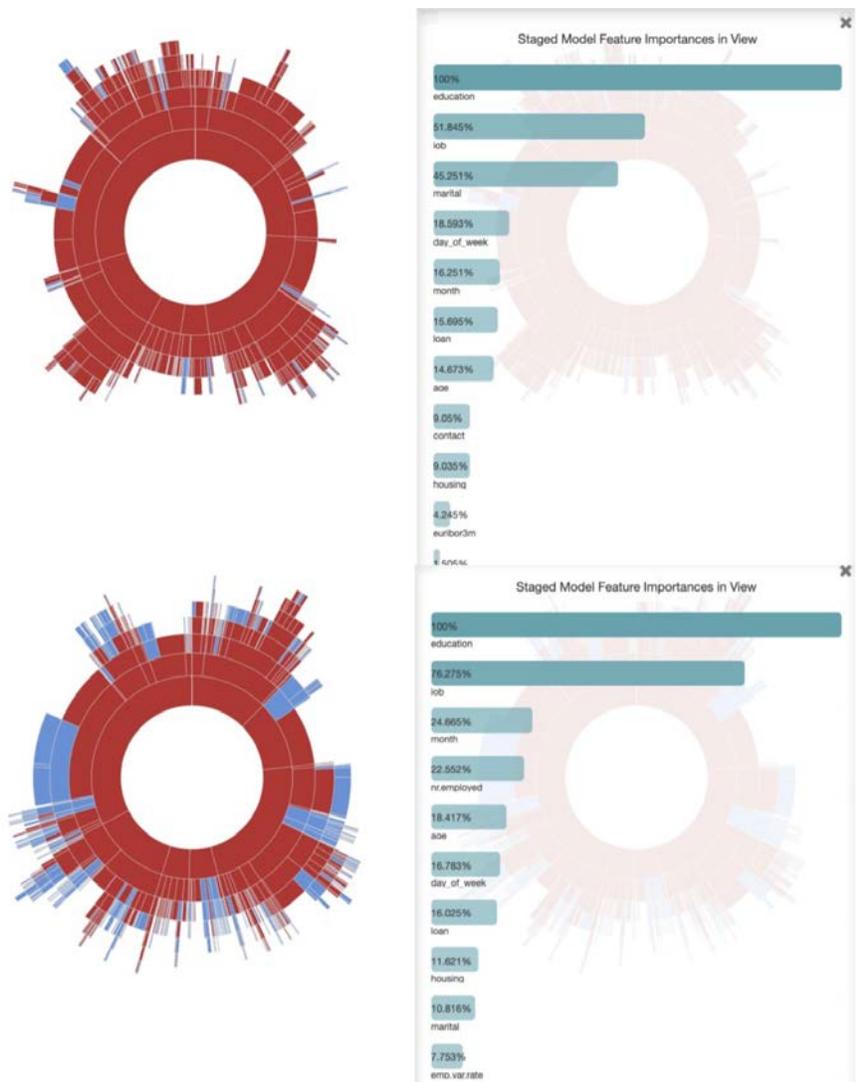


Figure 5: Bank Marketing Data, Explanation models for Batch Machine Learning (top) and Continuous Machine Learning (bottom).



We would like to note that there are approaches that automatically detect concept drift and decide to train a model when enough change happens [Velipasaoglu, 2018]. When continuous learning, by default, we accept that change happens all the time and hence go along and train all the time. It should also be noted here that, batch learning with small periods and continuous learning are not the same. The accuracy with small enough periods can be quite close to that of continuous learning. But the computing costs increase vastly as training becomes more frequent. Especially as we use machine learning more, it is important that we behave responsibly in terms of our energy consumption and damage to the environment. Also there is the cost of deploying models again and again. If tens or hundreds of deployed mini-batch machine learning models are used, if they all change, there would be a costly retraining, redeployment and adjustment phase.

Please note that for our experiments on the Bank Marketing dataset, in order to prevent data leakage, the “duration” feature was ignored. If the duration of the call is known, then the call is already placed anyway.

References:

[Barros2018] Barros, R. S. M., & Santos, S. G. T. C. (2018). A large-scale comparison of concept drift detectors. *Information Sciences*, 451, 348-370.

[Gama2004] Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, 46(4), 44.

[Moro2014] Moro, S., Cortez, P., & Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62, 22-31. Dataset at: <http://archive.ics.uci.edu/ml/datasets/Bank+Marketing>

[Velipasaoglu2018] Velipasaoglu, E., Concept Drift: Monitoring Model Quality in Streaming Machine Learning Applications, <https://www.youtube.com/watch?v=woRmeGvOaz4>

[Zliobaite2016] Žliobaitė, I., Pechenizkiy, M., & Gama, J. (2016). An overview of concept drift applications. In *Big data analysis: new algorithms for a new society* (pp. 91-114). Springer, Cham.

Tazi.ai

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



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.ai

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



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