Designing for Trust with Machine Learning

Fabien Girardin, Pablo Fleurquin

BBVA Data & Analytics, Avenida de Burgos, 16D 28036 Madrid, Spain  
fabien.girardin@bbvadata.com, pablo.fleurquin@bbvadata.com

Abstract

This is a proposal for a presentation at the Designing the User Experience of Machine Learning Systems symposium as part of the 2018 AAAI Spring Symposium Series in Palo Alto, CA. Trust is at the bedrock of our human social system. Historically, the financial industry has been based on how it could trust customers, and not the other way around. Today, customers request — in addition to competence, security and lending capability — fairness, honesty, legibility, transparency and other key attributes of the trust relationship with their bank. We will share learnings from our experiments that use Machine Learning techniques to design for trust.

Introduction*[[1]](#footnote-1)*

Fabien Girardin is Co-CEO at BBVA Data & Analytics, a center of excellence in financial data analysis than transforms the banking industry in the domains of marketing intelligence, customer advisory, risk, fraud and the automation of financial processes. With a broad spectrum of interdisciplinary skills, he guides teams in transforming algorithmic research and experiments into implications, value propositions, services, products and experiences that are future forward.

Pablo Fleurquin is Data Scientist at BBVA Data & Analytics with extensive experience in describing, analyzing and modelling the delay dynamics of a paradigmatic socio-technical complex system such as the air-transportation system. He uses his knowledge in Complex Network Theory, Graph Analytics and Machine Learning to develop online credit card fraud analytics, risk scoring solutions and pricing strategies.

This paper reports on our investigation and experiments that explore on how the specific design of Machine Learning algorithms can consolidate trust in financial services. We aim this work to help orient the design for automation and trust of any data-driven organization.

An Evolution of Trust

Trust is part of a social contract with rational and emotional bonds. Trust cannot be delivered, but actions can be taken in order to enrich it. For instance, financial businesses are based on how it could trust customers, and not the other way around.

Historically, in the financial industry, quality, transparency and altruism was demanded on the side of the customer. In consequence, a bank is often perceived as a partner people need to live with, but that are prone to mislead, provoke unfair situations and take advantage of opaque processes. That situation is changing with regulators and society, in various parts of the world, demanding openness for both protection of personal data and breaking bank’s monopoly in handling user data. Customers request — in addition to competence, security and lending capability — honesty, legibility, transparency and other key attributes of the trust relationship with their bank.

An Evolution of Automation

Many contemporary banks follow a digital transformation focused on automating sales, so more personalized offers reach customers at the right moment, and automating servicing though ‘Do it Yourself’ experiences, which allows for huge cost reductions on mature, high margin products. They employ Machine Learning to perform more fine-grained assessment of risks and provide customized advisory. Consequently, the trust relationship becomes more symmetric, because banks come to customers rather than customer go to the. We are investigating how the design of automation with Machine Learning techniques can actually enrich trust.

Experiments on Trust and Machine Learning

The following experiments aim at understanding techniques to design for trust with Machine Learning

* Interpretability and trust in credit risk scoring: Algorithmic transparency is openness about the purpose, structure and underlying actions of the algorithms used to search for, process and decision making. This experiment explores one way of making a black-box algorithm transparent using LIME as an interpretability framework.
* Learning to bid in real time using a fair strategy: this experiment evaluates an approach on dynamic pricing that uses Reinforcement Learning (RL) to keep a balance between revenue and fairness. This type of algorithm helps maximize revenues while taking into account fairness and equity that prevent a negative user perception of unfair price differences that can destroy a trustful relation.
* A fair approach on Recommender Systems (RS): While RS aim to provide an appealing list of items to users, most algorithms suffer from a bias in the recommendation towards popular items. As a consequence, the recommended list often goes away from the true interest of users. On the other hand, less popular, long-tail items are desirable for recommendations because of their novel and diverse character. In this experiment, we explore the concept of fairness in recommender systems, so that all items have the same chance to be presented to users.

Conclusion

So far, technology supported banks’ primary focus of as profit seeking machines rather than trust building machines. With the new wave of Machine Learning solutions, value is created with an accumulation of touch points that feed algorithms with behavioural data. Trust is enriched through the intensifying relations, feedback loops, virtuous cycles, ‘data network effects’, and the capacity to understand and react on customer’s intentions, emotions, and behaviours. Therefore, we believe the automation of the relation between people and services requests a specific design for trust.

1. Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved. [↑](#footnote-ref-1)