

Machine Learning for Q&A sites: state of the art and research directions

Xavier Amatriain
Quora Inc.
650 Castro St #450
Mountain View, CA, US
xavier@quora.com

ABSTRACT

Q&A sites like Quora aim at growing the world's knowledge. In order to do this, they need to get the right questions to the right people to answer them, but also the existing answers to people who are interested in them. In order to accomplish this they need to build a complex ecosystem issues such as content quality, engagement, demand, interests, or reputation are taken into account. It is not possible to build a system like this unless most of the process are highly automated and scalable. The good news is that using high-quality data you can build machine learning solutions that can help address all of the previous requirements.

In this talk I will describe some interesting uses of machine learning for Q&A that range from different recommendation approaches such as personalized ranking to classifiers built to detect duplicate questions or spam. I will describe some of the modeling and feature engineering approaches that go into building these systems. I will also share some of the challenges faced when building such a large-scale knowledge base of human-generated knowledge. Finally, I will describe some of the unresolved research challenges in the Q&A Space.

I will use my experience at Quora¹ as the main driving example. Quora is a Q&A site that despite having over 80 million unique visitors a month, it is known for keeping a high-quality of answers and content in general.



1. BIO

Xavier Amatriain is VP of Engineering at Quora, where he leads the team building the best source of knowledge in the Internet. With over 50 publications in different fields, Xavier is best known for his work on Machine Learning in general, and Recommender Systems in particular. Before Quora, he was Research/Engineering Director at Netflix, where he led the team building the famous Netflix Recommendation algorithms. Previously, Xavier was also Research Scientist at Telefonica Research and Research Director at UCSB. He has also lectured at different universities both in the US and Spain and is frequently invited as a speaker at conferences and companies.

¹<http://www.quora.com>