Machine Learning Architecture - Part II

Matheus Morroni

January 2020

1 The Importance of a Good Architecture

When an enterprise is going to apply machine learning for its business, the IT team should develop a solid architecture. A good architecture covers all crucial issues likewise, business issues, data issues, security and privacy issues. And of course a solid architecture should deal with technical issues in order to minimize the risk of instant project failure.

Architecture is a minefield. And creating a solid architecture for an innovative ML system and application is an unmarked road. Architecture is not by definition high level and sometimes relevant details are not documented. But getting the details of the inner working on the implementation level of ML algorithms may be very hard. So the existence of a map of architecture on machine learning should help in many ways.

Unfortunately do not exist a single machine learning reference architecture. Architecture organizations are never the first runners with new technologies. So there are not yet many mature ML reference architectures. It is possible to find vendor specific architecture blueprints, but these architectures mostly lack specific architecture areas as business processes needed and data architecture needed. Also, these vendors architecture blueprints tend to sell a specific solution. And a specific solution is not always built to solve the problem of the company.

2 Examples of Machine Learning Features for Business

2.1 Collaborate

The successful creation of ML applications requires the collaboration of people with different expertises. The company needs, e.g. business experts, infrastructure engineers, data engineers and innovation experts. The implication of this feature is the organisational and culture must allow open collaboration.

2.2 Unfair Bias

The ML algorithms and datasets can reflect, reinforce, or reduce unfair biases. Recognize fair from unfair biases is not simple, and differs across cultures and societies. However, always make sure to avoid unjust impacts on sensitive characteristics such likewise, race, ethnicity, gender, nationality, income, sexual orientation, ability, and political or religious belief. The implication of this feature is the transparency about the data of the company and training data sets, make models reproducible and auditable.

2.3 Built and Test for Safety

Use safety and security practices to avoid unintended results that create risks of harm. The company should design machine learning driven systems to be appropriately cautious. The implication of this feature is the perform risk assessments and safety tests.

2.4 Privacy by Design

The privacy by principles is more than being compliant with legal constraints. It means that privacy safeguards, transparency and control over the use of data should be taken into account from the start. This is a hard and complex challenge.

2.5 Contraints about Machine Learning Architecture for Business

The most important issues for a machine learning reference architecture are the aspects:

- Business aspects (e.g. capabilities, processes, legal aspects, risk management);
- Information aspects (data gathering and processing, data processes needed);
- Machine learning applications and frameworks needed (e.g. type of algorithm, an interface with a good UX);
- Hosting (e.g. compute, storage, network requirements but also container solutions);
- Security, privacy and safety aspects;
- Maintenance (e.g. logging, version control, deployment, scheduling);
- Scalability, flexibility and performance.

3 References

- Machine Learning Architecture, educba.com
- ML Reference Architecture, freeandopenmachinelearning.readthedocs.io/en/latest/architecture