Preface

These proceedings contain the 4 papers accepted to, and presented at, the research track of the 2nd International Conference on Predictive APIs and Apps (PAPIs '15), held in Sydney, Australia on August 7th, 2015.

Predictive Application Programming Interfaces (APIs) are receiving a lot of interest in the industry as they accelerate the development of predictive apps, by making it easier for application developers to use predictive models in production settings. They are a means of exposing predictive models to other programs, and they can also expose model learning capability, in which case one may speak of Machine Learning (ML) APIs. They exist in commercial offerings, such as Microsoft Azure ML and BigML presented at the research track, and Amazon ML, Datagami and Google Prediction API presented during the industry track, where ML algorithms run on cloud platforms and are accessed "as a service" (MLaaS). Predictive APIs can also be created from open-source or custom frameworks and self-hosted, as demonstrated in the Upwork and PSI presentations of the research track, but also in the industry track with Seldon and PredictionIO (at PAPIs '14).

MLaaS seeks to abstract away most of the complexity of building and running predictive models at scale, thus making ML easier and quicker to deploy, but also more accessible to developers. In general, exposing models as REST (http) APIs is a way to "package" them so that they can easily be used by desktop, web or mobile developers—who may be using different programming languages and technologies than those used by ML practitioners to build models (e.g. Python, R, Scala, Octave).

The first paper of these proceedings gives us a behind-the-scenes look at Microsoft Azure ML, an MLaaS environment for authoring predictive models, experimenting with them, running them on a cloud infrastructure and publishing them as web APIs. The Azure ML team presents design principles, challenges encountered and lessons learnt while building the platform.

While it is common for ML practitioners to measure models' performance via predictions' accuracy, the second paper of these proceedings by Brian Gawalt of Upwork focuses on concerns of software engineers who are in charge of deploying in production and scalability: models' throughput and response time. Using today's MLaaS products can help on those aspects, as they scale automatically up to a certain point. But for applications that require low latency and very high throughput (e.g. web services with millions of users), Gawalt offers an open-source API solution that parallelises data retrieval, featurization, learning and predictions, based on the Actor framework of concurrent programming. His solution provides a mid-point between from-scratch concurrent frameworks and mega-scale industry offerings (e.g. Spark), which is particularly suitable for lone data scientists looking to significantly increase predictions throughput.

In the third paper, Montgomery et al. of the PSI project introduce an open standard for REST Machine Learning APIs. One main advantage of exposing training and usage of predictive models via APIs is separation of concerns, but for these APIs to reach their full potential, they need to be easily interchangeable and interoperable. This would allow app developers, for instance, to test different APIs by just changing their URIs in their code, and to use combinations of the best of the MLaaS and open-source worlds to create complex predictive features in their apps.

Finally, the last paper of these proceedings by Poul Petersen of BigML re-situates Machine Learning APIs in the history of the field, reviews practical challenges in ML usage in industry, identifies new trends in cloud-based machine learning APIs (such as composability and specialization) and outlines the company's vision for the future of these APIs.

Standards, interoperability and composability were also mentioned in Robert Williamson's invited keynote at the conference, and in the panel discussion on research challenges facing Predictive APIs. Other topics discussed were privacy and security, accountability and transparency, automatism (in model selection, detecting data types, data wrangling, text featurization), trust in APIs' models, and alignment with business concerns. We expect to hear more about these topics at future editions of PAPIs!

Acknowledgments

Program Committee: Erick Alphonse, Sébastien Arnaud, Danny Bickson, Misha Bilenko, Christophe Bourguignat, Mikio Braun, Natalino Busa, Sharat Chikkerur, Matthew Grover, Harlan Harris, Jeroen Janssens, Jacek Kowalski, Nicolas Kruchten, Greg Lamp, Gideon Mann, Francisco Martin, Joe McCarthy, Charles Parker, Mark Reid, Thomas Stone, Keiran Thompson, Slater Victoroff, Zygmunt Zajac. **Student Volunteers:** Johann Bengua, Anantasri Boddupalli, Nazanin Borhan, Joan Capdevila Pujol, Alex Cullen, Min Fu, Yannis Ghazouani, Shameek Ghosh, Asso Hamzehei, Andrew Lee, Rongzuo Liu, Jose Magana, Belinda Qin.

In addition to the entire program committee and student volunteers, we would like to thank the Local Chair Gerry Carcour, who took care of local coordination and arrangements for PAPIs '15, and the Media and Publicity Chair Ali Syed. We are grateful for the financial support provided by our corporate sponsors BigML, NVIDIA, GCS Agile and Datagami. Finally, we wish to thank the authors for their time and effort in creating such a fine program.

February 2016

The Editorial Team:

Louis Dorard, PAPIs '15 General Chair
louis@dorard.me

Mark D. Reid, PAPIs '15 Research Chair
Australian National University
mark.reid@anu.edu.au

Francisco J. Martin, PAPIs '15 Program Chair
BigML
martin@bigml.com

Table of Contents

Introduction