A PROPOSED MODEL ARTIFICIAL INTELLIGENCE GOVERNANCE FRAMEWORK

January 2019

TABLE OF CONTENTS

FOREWORD

From the well-publicised achievements of Google's DeepMind, SenseTime's technologies on facial recognition, to the ubiquitous presence of virtual assistants like Apple's *Siri* or Amazon's *Alexa*, Artificial Intelligence ("AI") is now a growing part of our lives. AI has delivered many benefits, from saving time to diagnosing hitherto unknown medical conditions, but it has also been accompanied by new concerns such as over personal privacy and algorithmic biases.

Amid such rapid technological advances and evolutions in business models, policy makers and regulators must embrace innovation in equal measure. The genesis of this Model AI Governance Framework ("Model Framework") can be traced to efforts by policy makers and regulators in Singapore to articulate a common AI governance approach and a set of consistent definitions and principles relating to the responsible use of AI, so as to provide greater certainty to industry players and promote the adoption of AI while ensuring that regulatory imperatives are met. This Model Framework is adapted from a discussion paper issued by the Personal Data Protection Commission (PDPC) in June 2018.

The first edition of this accountability-based Model Framework aims to frame the discussions around the challenges and possible solutions to harnessing AI in a responsible way. The Model Framework aims to collect a set of principles, organise them around key unifying themes, and compile them into an easily understandable and applicable structure. It seeks to equip its user with the tools to anticipate and eventually overcome these potential challenges in a practical way.

The Model Framework is Singapore's attempt to contribute to the global discussion on the ethics of AI by providing a framework that helps translate ethical principles into pragmatic measures that businesses can adopt. The Model Framework has been developed in consultation with academics, industry leaders and technologists from different backgrounds and jurisdictions. This diversity of views reflects the desire of the PDPC, the Info-communications Media Development Authority (IMDA), and the Advisory Council on the Ethical Use of AI and Data, to shape plans for Singapore's AI ecosystem in a collaborative and inclusive manner.

Where AI is concerned, there are big questions to be answered, and even bigger ones yet to be asked. The Model Framework may not have all the answers, but it represents a firm start and provides an opportunity for all – individuals and organisations alike – to grapple with fundamental ideas and practices that may prove to be key in determining the development of AI in the years to come.

S Iswaran
Minister for Communication and Information
Singapore
January 2019

ANNEX C

Use Case in Healthcare – UCARE.AI

UCARE.AI (https://www.ucare.ai) is an artificial intelligence and machine learning company on a scientific mission to solve healthcare problems and advance humankind through the ethical and responsible use of data. UCARE.AI deploys a suite of AI and machine learning algorithms, including proprietary deep learning and neural network algorithms, built on a cloud-based microservices architecture to provide sustainable and customisable healthcare solutions for doctors, hospitals, patients, insurers and pharmaceutical companies.

A successful use case is the recent implementation of AI-Powered Pre-Admission Cost of Hospitalization Estimation (APACHE™) for four major hospitals, namely Mount Elizabeth, Mount Elizabeth Novena, Gleneagles and Parkway East hospitals; owned by Parkway Pantai. This study shares UCARE.AI's methodology for developing and deploying APACHE, a scalable plug-and-play system that provides high availability, fault-tolerance, and real-time processing of high-volume estimate requests. APACHE provides more accurate estimates, with a four-fold improvement in accuracy over Parkway Pantai's previous bill estimation system. This is done with the intent of achieving standardisation of healthcare cost estimation and provision of greater price transparency to facilitate the building and maintenance of trust between payers, providers, and patients. This is in line with UCARE.AI's commitment to ensure patients continue to make well-informed decisions on available medical treatment options.

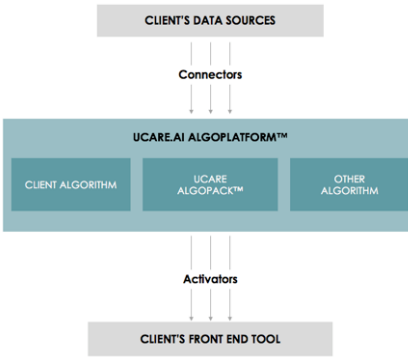
Background

Previous healthcare cost estimation methods involve traditional techniques such as (i) normal distribution-based techniques, (ii) parametric models based on skewed distributions, (iii) mixture models, (iv) survival analysis, etc. The existing approach used was via simple statistical aggregations based on the Table of Surgical Procedures quoted prices or ICD-10 diagnostic codes.

Challenges include relatively high error rates, high financial and human cost of updates, and low frequency of updates due to these high costs.

UCARE.AI worked with Parkway to resolve these issues with a multi-step process involving: (i) data exploration, (ii) data cleaning, (iii) feasibility assessment (iv) feature engineering, (v) machine learning, and (vi) presentation of results. With satisfactory results from the proof of concept, APACHE was then put into production.

High-Level Architecture of APACHE API



1. *Data Sources.* Relevant data is obtained from partner organisations for use. As the system is further improved upon, publicly available data sources as well as third-party data are used to generate predictions, thereby reducing the need for personal data collection.
2. *Connectors.* Basic data validation is conducted prior to being ingested into the data production warehouse.
3. *AlgoPlatform.* The data is processed by the algorithms, and encrypted for storage. The algorithms are integrated with reporting and monitoring systems for performance management and intervention to minimise downtime. Various machine learning models can be deployed to allow for model comparisons and can be hot-swapped in a live production environment.
4. *Activators.* These serve to assist with data authentication and verification, to send results to the client's chosen front end tool.

Aligning with PDPC's Model AI Governance Framework

UCARE.AI adopts a proactive approach that aligns with PDPC's Model AI Governance Framework.

Trustworthy and Verifiable

The proposed AI governing framework acknowledges that neural networks are inscrutable and verification of the results provided by such networks is required prior to putting them to use in human applications. UCARE.AI circumvents this problem by continuously validating the accuracy of its algorithms against the ground truth. Weekly check-ins with participating partners and domain experts are also employed to ensure quicker and more reliable iterations. Automated re-training of the data models ensure that the algorithms remain up-to-date. This methodology of continuous validation of its AI models with the help of experts from Parkway Pantai will help to boost confidence in the accuracy of its predictive insights and will help train algorithms to become even more precise with each amount of data inputted.

Accountability and Transparency

Prior to data collection, informed consent from stakeholders would have been obtained and approval of the use of data sought via open communication channels. The careful curating and conversion of data into usable format prior to building the models ensures the AI algorithm is kept accountable and coherent to users; this is done in conjunction with Parkway Pantai. The proper storage and repair of previously broken or missing data also serve to provide greater transparency and safety to users by minimising the influence of data gaps in the projection of the result. Careful monitoring of data is key in ensuring service reliability, and therefore detailed and consistent logging across the multiple components involved is also employed in APACHE, collected in a secure, centralised log storage that is made easily accessible to the development and operations team when required, allowing for prompt debugging and uptime tracking if necessary.

Fairness

The automated prediction of hospitalisation costs reduces the likelihood of human biases affecting the ultimate judgement of the data and provides an element of consistency across all predictions. Discrimination based on income levels and insurance coverage, for instance, would be effectively negated. Although there would be concerns about the use of a 'human-out-of-the-loop' system, the algorithm in question is designed to be human-centric.

Human-Centric

This use case highlights how artificial intelligence may be used in augmenting decision-making capabilities in a human-centric manner whilst minimising the potential risks of harm to involved parties. The automated process of bill estimation negates the need for tedious statistical calculations, thereby freeing up man-hours and effort to allow for the channeling of these into more creative pursuits. Furthermore, the information provided would serve to benefit patients and payers by allowing for more accurate cost forecasting, efficient allocation and distribution of healthcare resources, and guidance on new policy initiatives. Patients would be conferred greater peace of mind over their healthcare expenditure such that they may focus their energies on recovery instead.

To minimise the risk of harm, rigorous feasibility studies are conducted prior to using the data to focus on creating a valid and robust validation framework. This will be done in conjunction with partners and their feedback on the proposed framework obtained before proceeding. A human feedback loop with inputs from the client organisation (Parkway Pantai-owned hospitals) is also in-built into each algorithm to enhance sophistication, while a manual override protocol is also included to ensure that these algorithms can be safely terminated if deemed necessary. This ensures that the algorithm remains under human control and in line with the medical field's well-established ethical principles of beneficence, non-maleficence, and social justice.

For more information, please visit https://www.ucare.ai or contact hello@ucare.ai.

ACKNOWLEDGEMENTS

The Personal Data Protection Commission, Infocomm Media Development Authority, expresses its sincere appreciation to the following for their valuable feedback to this Model AI Governance Framework (in alphabetical order):

AIG Asia Pacific Insurance Pte. Ltd.
Asia Cloud Computing Association
AsiaDPO
BSA | The Software Alliance
DBS
Element AI
Fullerton Systems and Services
Grab
IBM Asia Pacific
MasterCard
Microsoft Asia
OCBC Bank
Salesforce
Standard Chartered Bank
Temasek International
Ucare.AI

END OF DOCUMENT