Cognitive Compliance: Assessing Regulatory Risk in Financial Advice Documents

Wanita Sherchan, Sue Ann Chen, Simon Harris, Nebula Alam,
Khoi-Nguyen Tran, Christopher J. Butler
IBM Research Australia
{wanita.sherchan, sachen, siharris, anebula, khndtran, chris.butler}@au1.ibm.com

Abstract

This paper describes Cognitive Compliance - a solution that automates the complex manual process of assessing regulatory compliance of personal financial advice. The solution uses natural language processing (NLP), machine learning and deep learning to characterise the regulatory risk status of personal financial advice documents with traffic light rating for various risk factors. This enables comprehensive coverage of the review and rapid identification of documents at high risk of non-compliance with government regulations.

1 Introduction

In response to the increasing demand for financial advisory services, recent years have seen the introduction of tighter government regulations to ensure that financial advisors act in the "best interests" of the client (FOF 2012). In Australia, approximately 5 to 15% of the financial advice documents (Statement of Advice, or SoA) produced are audited by the government (ASI 2018a). These manual and protracted audits found that 75% of the SoAs reviewed were non-compliant (ASI 2018b). In addition, the retrospective nature of the audits fail to protect customers from financial losses which may result from poor advice.

Addressing these issues, we have developed Cognitive Compliance - a solution that leverages natural language processing, machine learning and deep learning methods to assess the compliance status of SoA documents characterised by traffic light ratings for various risk factors. The solution enables rapid review of all SoAs with respect to risk of non-compliance in several aspects. Advisors can perform the reviews before the SoA is formalised/executed to ensure they are compliant with the regulatory requirements; and consequently clients are protected from harmful advice. Further, the auditors are able to efficiently review a large number of advice documents to focus on those at high risk of non-compliance.

Existing commercial solutions such as Logical Construct[1] and ThoughtRiver[2] focus on assessing risk in contractual documents. Rubik[3] provides a financial planning software platform for generating advice documents. Advice RegTech[4] provides software for assisting licensees in interpreting best interest duty and related obligations and for assessing potential risk in their licensed advisors. In contrast, our solution focuses on assessing regulatory risk in advice documents and provides a traffic light summary of various key risk indicators in an SoA. Moreover, our solution supports and benefits clients, advisors and auditors alike.

2 Risk Factors Modeling

We identified a set of key risk indicators (KRIs) affecting the compliance status of SoA documents in a study with an Australian government regulatory agency and Promontory, a global consulting firm specialising in regulatory and risk management advice to financial institutions (Sherchan et al. 2019). We then developed a number of machine learning, deep learning and rules-based models to extract, score and validate these risk indicators on a diverse collection of 400 anonymised SoA documents from various financial institutions. We briefly describe our approach for extracting and scoring each of the KRIs below.

1. Goal-advice mapping: We employ a Long Short-Term Memory (LSTM) model to extract sentences that describe a client's goals, and the advisor's recommendations in the SoA. We then use a textual entailment method to assess the appropriateness of the advice to the goals (Chen et al. 2019).

2. Investment asset classes: We apply a random forest (RF) classifier to identify tables within the SoA that relate to asset classes, projections and cash flow analysis. This risk factor is satisfied if the SoA contains asset class tables.

3. Client's position after advice: We apply a Bag of Embeddings (BoE) model to extract sentences in the SoA that discuss the client's capital position. We then augment this with the projections table identified using the above-mentioned RF table classifier to determine the risk rating.

4. Cash flow analysis: To determine whether the SoA contains an analysis of the client's post-advice cash flow, we use the RF table classifier mentioned above. Following this, we

[1] Logical Construct: https://www.logicalconstruct.com/ [2] Thought River: https://thoughtriver.com

[3] Rubik: http://rubik.redc.me/products/coin [4] Advice RegTech: https://adviceregtech.com/

apply a set of rules to assess whether this analysis shows a positive outcome for the client.

5. Insurance consideration: We employ a BoE model to determine whether an SoA recommends personal insurance, defers insurance discussions or excludes insurance discussion from the SoA. A rules-based model is then applied to determine a risk rating for this indicator.

3 Solution Architecture and Deployment

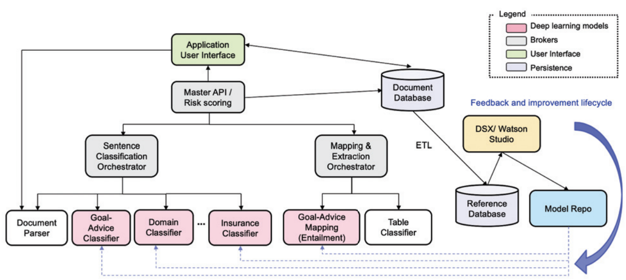

Figure 1: Cognitive Compliance solution architecture

Figure 1 shows the architecture of the Cognitive Compliance solution. The Artificial Intelligence (AI) models were developed using AllenNLP (Gardner et al. 2017), an open-source NLP Python library, and deployed using Flask[5] and Swagger[6]. The AI models and orchestrator services are deployed as microservices that are containerised in Docker[7]. This enables segregated build and runtime environments and deployment on multiple cloud platforms. As such, our solution is currently deployed in the IBM Cloud™ platform[8] and in an IBM Cloud Private Kubernetes®[9] environment. Containerised deployment enables the AI models to be retrained and updated independently, an added advantage when incorporating new ground truth data obtained from user interactions with the solution. Furthermore, since Kubernetes supports automated roll-out of updated services, AI models can be redeployed without service interruption.

Initially the Document Parser ingests and parses each SoA document, after which the pipeline of KRI models are applied to each sentence in the document. The master API/Risk Scorer aggregates the model results at the SoA level and calculates the risk rating per KRI. The result is traffic light annotation of the document for each KRI which is then presented in the user interface. The solution uses CouchDB[10] as the JSON document store. CouchDB offers versioning of documents, a useful feature when updating and re-uploading SoA documents, and when user interaction with the application changes the document. A web application built using ReactJS[11] framework on top of a NodeJS[12] server provides the user interface for the solution allowing the user to easily identify those SoAs at highest risk of non-compliance and drill down into individual KRIs within those SoAs. Figure 2 gives an overview of an SoA as presented by our system.

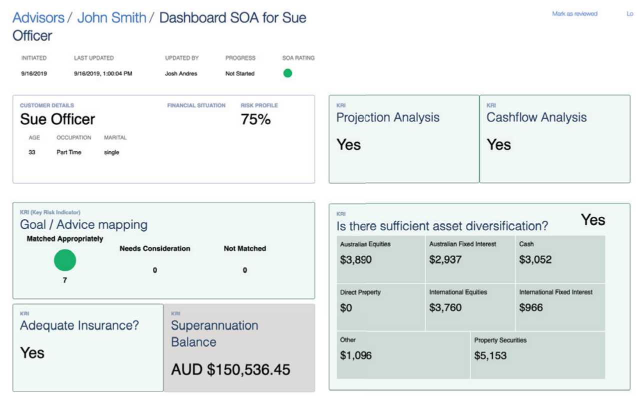

Figure 2: Overview of an SoA with traffic light scoring of the Key Risk Indicators.

All user interaction is captured by the system for auditing purposes and for further refining the AI models.

4 Conclusion

In this paper we have described Cognitive Compliance - a novel solution that applies natural language processing, machine learning and deep learning techniques to extract and score key risk indicators impacting the regulatory compliance status of financial advice documents. This solution enables bulk analysis of SoAs by reducing the time taken for auditing advice documents from several hours to minutes, highlighting the most at-risk documents and identifying the sources of those risks; making it an essential tool for assessing regulatory compliance of financial advice documents.

References

2018a. Report 515 Financial advice: Review of how large institutions oversee their advisers. Technical report, Australian Securities and Investments Commission, Australia.

2018b. Report 562 Financial advice: Vertically integrated institutions and conflicts of interest. Technical report, Australian Securities and Investments Commission, Australia.

Chen, S. A.; Makarucha, A. J.; Alam, N.; Sherchan, W.; Harris, S.; Yiapanis, G.; and Butler, C. J. 2019. Evaluating goal-advice appropriateness for personal financial advice. In *IEEE Global Conference on Signal Information and Processing*.

2012. Corporations Amendment (Future of Financial Advice Measures) Act 2012. Federal Register of Legislation.

Gardner, M.; Grus, J.; Neumann, M.; Tafjord, O.; Dasigi, P.; Liu, N. F.; Peters, M.; Schmitz, M.; and Zettlemoyer, L. S. 2017. Allennlp: A deep semantic natural language processing platform.

Sherchan, W.; Harris, S.; Chen, S. A.; Alam, N.; Tran, K.-N.; Makarucha, A. J.; and Butler, C. J. 2019. Assessing regulatory risk in personal financial advice documents: a pilot study. In *AAAI Fall Symposium on Artificial Intelligence in Government and Public Sector (AAAI/FSS-19)*.

[5] http://flask.pocoo.org/ [6] https://swagger.io/
[7] https://www.docker.com/ [8] https://cloud.ibm.com/
[9] https://kubernetes.io/ [10] http://couchdb.apache.org/
[11] https://reactjs.org/ [12] https://nodejs.org/