Predicting Decisions in the Supreme Court of the Philippines Using a Natural Language Processing and Machine Learning-based Framework *

John Kevin Abonita, Jeffrey A. Aborot, Roxanne Avinante, Rother Jay Copino
Michelle P. Neverida, Vanessa O. Osiana, Elmer C. Peramo, Glenn Brian Tan
Michael Virtucio and Joanna G. Syjuco
Research and Development Group
Computer Software Division
Advanced Science and Technology Institute
Department of Science and Technology
email: jep@asti.dost.gov.ph

Abstract

AAAI creates proceedings, working notes, and technical reports directly from electronic source furnished by the authors. To ensure that all papers in the publication have a uniform appearance, authors must adhere to the following instructions.

Introduction

The Philippine court system is faced with a large volume of backlog cases. The volume of this backlog continue to grow every year despite the successful projects of the Supreme Court aimed towards increasing the efficiency of the judges and their support staff

cite reference

. The court system lacks enough judges and court staff to accommodate the need for judicial services in the country. The country has only 2,000 courts available to serve the more than 1 million Filipinos nationwide. This averages to 1 court per 50,000 people Ref: http://asiafoundation.org/2016/12/14/conversation-philippines-chief-justice-maria-lourdes-sereno/. Specifically in the lower courts, the average number of cases expected to be handled is around 1 million cases per year. This averages to about 4,000 cases per court per day and 644 cases per judge per year (Ref: http://nap.psa.gov.ph/beyondthenumbers/2013/06132013_jrga_courts.asp). Case backlogs in courts also result to overcapacity of detention cells. Suspects in a case are put into detention until they are tried or until they pay bail for conditional release. Suspects with low income on the other hand are unable to pay bail and opt to be detained until their trial instead. Pre-trial detainees compose the 64% of prison population in the country (Ref: http://www.prisonstudies.org/country/philippines). Case backlogs in the Philippine courts has a far reaching effect in the lives of the citizens and this problem is evident mainly on the front-line courts, closer down to the people

cite reference

. The first-level (Metropolitan Trial Courts in Cities, Municipal Trial Courts, Municipal Circuit Trial Courts and Shari'a Courts) and second-level (Regional Trial Courts) lower courts suffer the most from case backlogs

cite reference

. High-impact solutions to the case backlog problem then need to be focused on these lower level courts. Specifically, focus must be put into the rate of delivery of decision on cases on the lower courts. The one clear way to dissipate backlog in courts is to raise the rate of delivery of decisions on cases higher than the rate of incoming new cases. This task is difficult for a judge who in average needs to dispense decision on two cases per day in a year.

The case backlog problem in the lower courts and the focus on increasing the rate of disposition of cases can be approached from various angles. Some of the reforms and projects of the Judiciary which contributed in the dissipation of the case backlog problem in the courts are the *Enhanced Justice on Wheels*, *Small Claims Courts*, *Zero Backlog Project*, *Hustisyeah!*

cite Judiciary annual report for 2013, 2015 and 2016

. Varying significant decrease in the case backlog of identified highly congested lower level courts resulted from these reforms

cite Judiciary annual report for 2013

. In this paper we look into the problem specifically from the technological perspective.

discuss briefly about NLP and ML specifically in application to legal documents

Nowadays, various Natural Language Processing (NLP) technologies coupled with Machine Learning (ML) technologies are already being used in different areas of the legal system. These application areas include *legal research*, *electronic discovery of evidences*, *contract analysis*, *legal education and legal prediction*

cite references for each area

. Specifically in the area of legal prediction, mathematical models are developed and used to predict outcomes of legal cases in courts

cite references for legal prediction

*This work is fully supported by the Advanced Science and Technology Institute of the Department of Science and Technology, Republic of the Philippines.

. The long-term aim of researchers in this area is not to replace human judges but to augment them in their decision-making process by providing initial classifications of cases based on *juris prudence*.

Specifically, we make the following contributions in this study

1. We provide pre-processed, cleansed and classified data sets of Philippine Supreme Court *juris prudence* cases from year 1987 to 2017 ready for use of the public for research purposes (see Data Acquisition and Cleansing).
2. We provide n-gram data sets resulting from natural language processing of the pre-processed Philippine Supreme Court *juris prudence* cases from year 1987 to 2017 ready for use of the public for research purposes (see Feature Extraction).
3. We provide various machine learning models for classifying future Philippine Supreme Court cases (see Model Development).

The Challenge

Our Solution and Its Technical Details

Theoretical Framework

Methodology

Results

Data Acquisition and Cleansing

Feature Extraction

Model Development Section

Discussion

Related Works

Conclusion

Acknowledgement

References