



Improved Scholarly Communication Using Machine Learning

Lighton Phiri | University of Zambia

Introduction

The University of Zambia is a public research university in Zambia. While there are numerous reported research activities at the institution, measuring the impact of this research is problematic due to the low online visibility of research output.

Challenges

- Research output not properly curated and archived.
- Discoverability of research output is compromised.
- Measuring the impact of research is difficult.

Current and Future Work

- Automate manual processes associated with curation of scholarly research output.
- Reclassification of curated digital objects.
- Implement toolkits for facilitating effective curation of scholarly research output.
- Bibliometric analyses of scholarly research output.

<oai_dc:dc>

<dc:title>

Evaluation of research education networks

</dc:title>

<dc:creator>Mwiinga, Jervas</dc:creator>

<dc:subject>

High performance computing--Zambia

</dc:subject>

<dc:subject>

Research education networks--Zambia

</dc:subject>

<dc:description>

[...]

</dc:description>

<dc:date>2018-07-23T13:00:50Z</dc:date>

<dc:date>2017</dc:date>

<dc:type>Thesis</dc:type>

[...]

<dc:format>application/pdf</dc:format>

<dc:publisher>

The University of Zambia

</dc:publisher>

</oai_dc:dc>

Preliminary Results

Preliminary work has focused on exploring techniques for automatically classifying Electronic Theses and Dissertations (ETDs), using features extracted from the single source of truth: ETD manuscript.

- Models for automatically classifying ETD collections, types and subjects implemented.
- Models deployed as a Flask-based API.

Figure 1. ETD Domain Classification

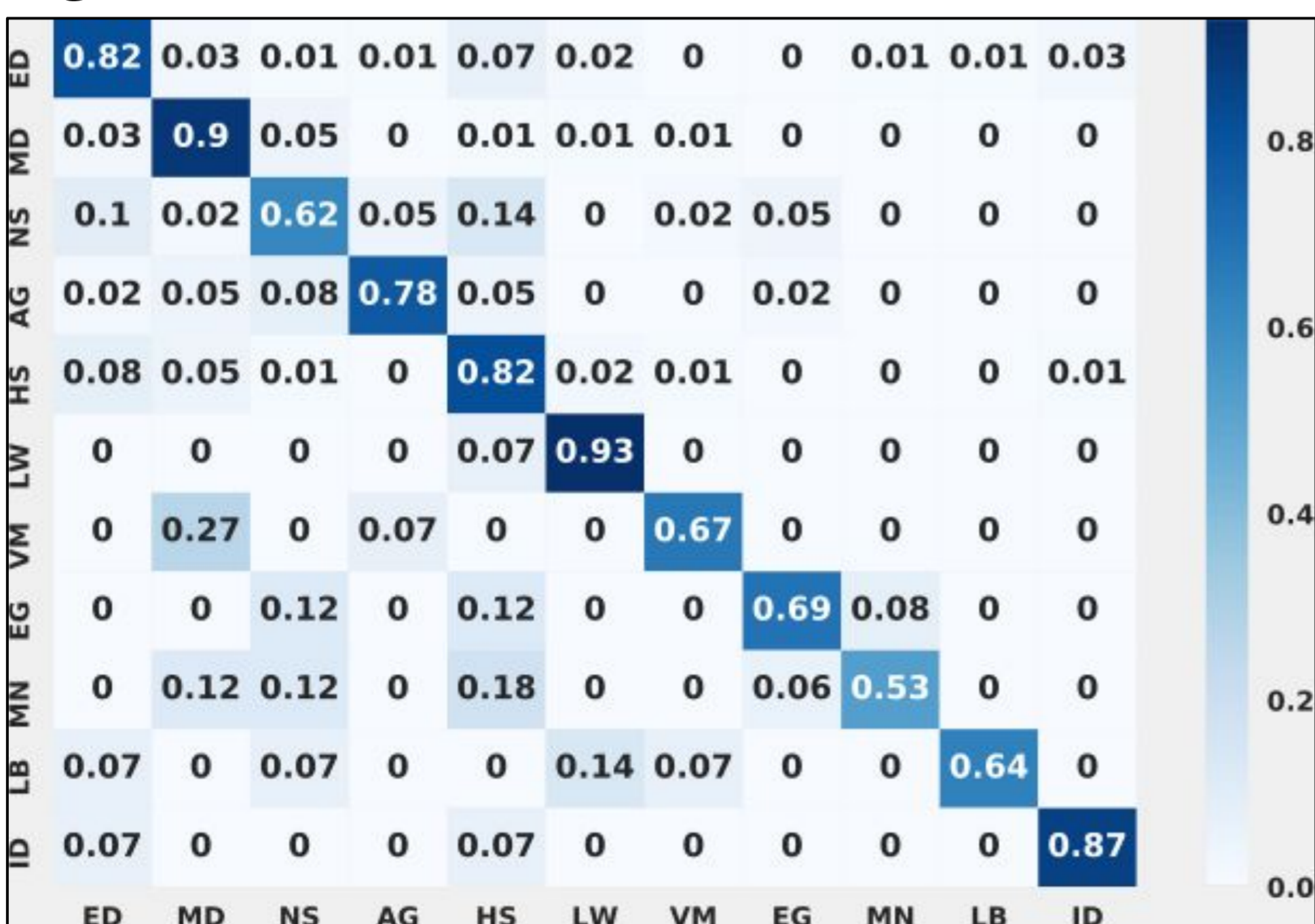
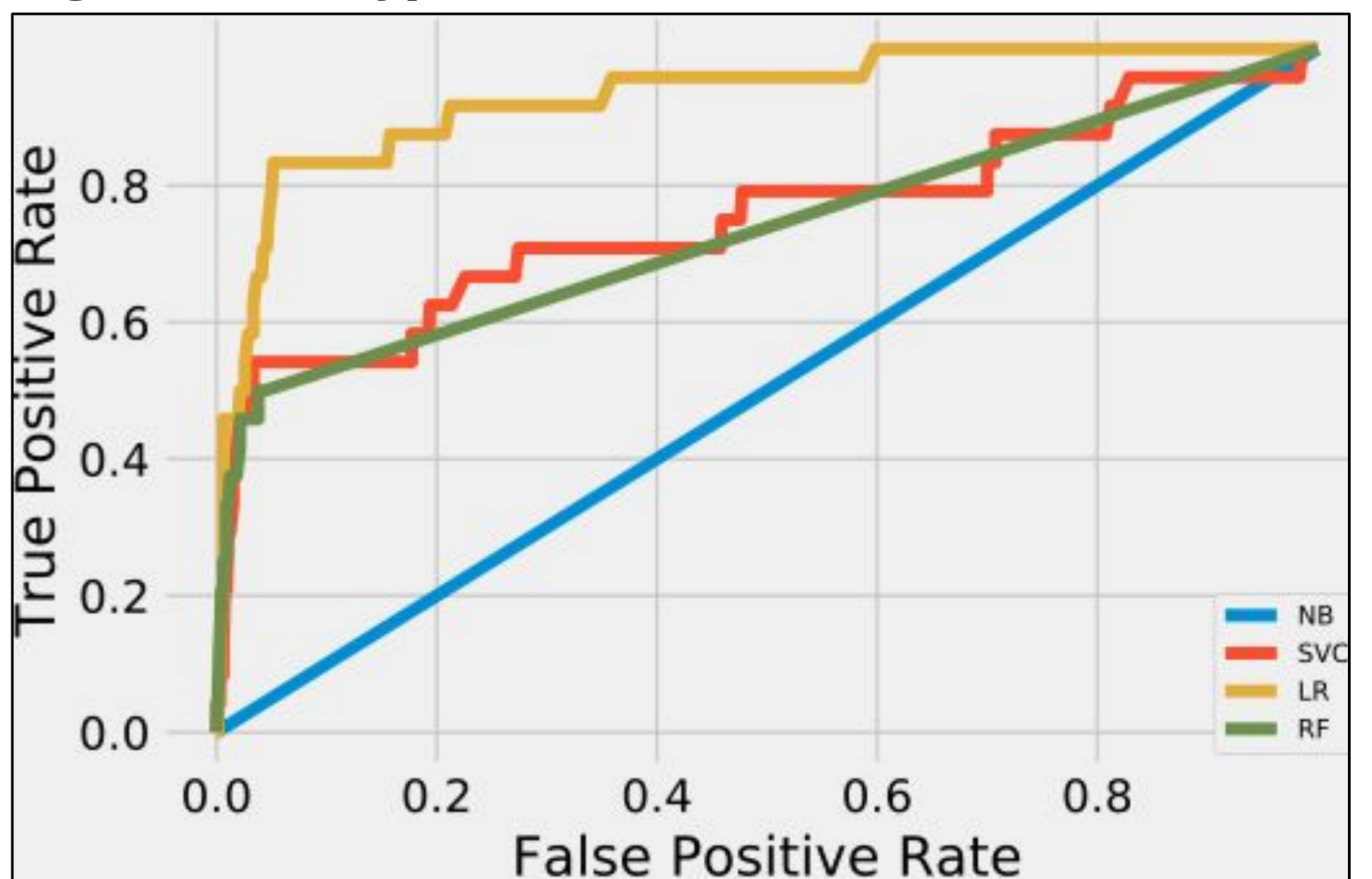


Figure 2. ETD Type Classification



ETD Classifier API: https://github.com/lightonphiri/etd_autoclassifier

