Analytical Support of Financial Footnotes Analysis

XBRL Conference
Maryam Heidari
Maryam.heidari@bwl.tu-freiberg.de
01.06.2016
Agenda

- Introduction
- Research Problem
- Research goal
- Research Methodologies
- Implementations
- Evaluation and Results
- Limitation and further Research
Introduction

Financial analysis Process (Bauwman et al., 1987)

Examination:
Establish the financial objective, select relevant and necessary data from different sources.

Integration and Observation:
Process data and calculating required ratios and trends.

Reasoning:
Interpret the results and make the financial decision.
Footnotes Tagging procedures (4 level)

Level 1- Block tagging

Level 2- Block tagging of significant accounting policies footnote

Level 3- Tables block tagging

Level 4- Detailed footnotes tagging
Introduction: Financial footnotes’ formats

- Financial footnotes
  - Traditional Footnotes
    - Paper, Pdf, or HTML
  - XBRL Footnotes
    - Block tagging footnotes (level 1-3)
    - Detailed tagging footnotes (level 4)
Research Problem

- Based on studied use cases
- Based on literature review:
  1. Financial analysis domain
  2. Using Text mining in financial domain
- According to expert interviews
  - Footnotes information can not be ignored.
  - No Automatic solution to extract/analysis
  - Extreme manual effort is needed
Research aims to design and evaluation of a practically applicable method to facilitate financial footnotes extraction/analysis process.
Methodology: Design Science Research (DSR) (Hevner, 2004)

Problem Identification
- Expert interview and literature review

Design and develop a prototype
- Text-classification prototype

Evaluation
- Experiment: human (expert) judges
Design and Implementation
Design and Implementation

- Tool: **Rapid miner**
- Database: **Edgar online database**
- Examined footnote: **Income tax** (6 categories are defined):

1. Deferred tax asset;
2. Effective tax rate;
3. Net operating loss (NOL);
4. Taxation authority;
5. Unrecognized tax benefit;
6. Valuation allowance

- Different classification algorithms (supervised and unsupervised) are applied.
- Different criteria are observed: accuracy, run time, ...
- The best results belongs to **Naïve Bayes** algorithm
# The Results of Text Classification Algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Run time</th>
<th>Accuracy</th>
<th>Absolute error</th>
<th>Precision</th>
<th>Recall</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-NN</td>
<td>7s</td>
<td>81.82%</td>
<td>0.183</td>
<td>82.34%</td>
<td>80.92%</td>
<td>0.362</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>4s</td>
<td>82.86%</td>
<td>0.171</td>
<td>78.48%</td>
<td>77.85%</td>
<td>0.414</td>
</tr>
<tr>
<td>SVM</td>
<td>28s</td>
<td>79.22%</td>
<td>0.784</td>
<td>80.37%</td>
<td>78.46%</td>
<td>0.786</td>
</tr>
<tr>
<td>Decision tree</td>
<td>1m 45s</td>
<td>90.65%</td>
<td>0.136</td>
<td>91.38%</td>
<td>90.82%</td>
<td>0.280</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unsupervised algorithm</th>
<th>K</th>
<th>Run time</th>
<th>Ave. within centroid distance</th>
<th>Davies-Bouldin</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-means</td>
<td>6</td>
<td>7min 20s</td>
<td>-0.827</td>
<td>-4.763</td>
</tr>
</tbody>
</table>
Qualitative evaluation: Experiment

**Goal:** to evaluate the efficiency of the text mining prototype on financial footnote analysis.

<table>
<thead>
<tr>
<th>Judges (experts):</th>
<th>ten financial experts in audit and financial advisory domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content of Experiment:</td>
<td>12 income tax footnotes from different industries</td>
</tr>
<tr>
<td>Minimum length:</td>
<td>3 pages</td>
</tr>
</tbody>
</table>
| Experts are given: | 1. Plain text of footnotes (input of the prototype)  
2. Classified footnotes (output of the prototype) |

The experts were asked to read each income tax footnote fully first and then to proceed to the related output table. Subsequently, they were invited to reply to the online questionnaire to rate the quality of the system based on different criteria using a five-point scale ranging from one to five.
Results and Discussion

<table>
<thead>
<tr>
<th>Relation between category and extracted sentence</th>
<th>Information access (time reduction)</th>
<th>Appropriate Information presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>77%</td>
<td>81%</td>
<td>85%</td>
</tr>
</tbody>
</table>
## Results and Discussion

<table>
<thead>
<tr>
<th>Different approaches to textual part of financial reports</th>
<th>Type of financial report</th>
<th>Description of method</th>
<th>Challenges and benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional approach</td>
<td>Traditional reports (Pdf, html, printed reports)</td>
<td>Manual process base on reading whole documents</td>
<td>Time consuming and error-prone</td>
</tr>
<tr>
<td>XBRL (block and detailed tagging)</td>
<td>XBRL reports</td>
<td>Tables and numbers are tagged but narrative information within footnotes lines should be analyzed manually</td>
<td>Both tagging procedures help just to have a unified structure in financial reports. Soft information are still read manually which it is Time consuming and error-prone</td>
</tr>
<tr>
<td>Content analysis based on coding system</td>
<td>Traditional reports (Pdf, html, printed reports)</td>
<td>Manual code assignment to the text</td>
<td>Codes should be assigned manually to the texts every time for each report. Mistakes of human coders. More accurate in comparison with manual processing</td>
</tr>
<tr>
<td>Proposed text classification prototype</td>
<td>Applicable for both traditional and XBRL reports</td>
<td>Automatic classification of financial footnotes based on pre-defined categories</td>
<td>Footnotes are classified automatically. Textual parts of footnotes are accessible through pre-defined categories. Time saving. More accurate in comparison with manual procedures. Training documents process have been done manually (for training the model). Increasing the number of categories reduces the accuracy of the prototype</td>
</tr>
</tbody>
</table>
Research Limitation

- Challenge of manual process of training documents
- Focused just on one of the footnote item (income tax). (limitation of the number of footnotes for prototype)
- Mapping financial footnotes into figures and values in financial statements.
- Expert interview and experiment phase conducted in German financial audit and financial advisory companies.
Theme 1: An additional study would be necessary to extend the model for more footnotes items such as leases, debt, and etc. An additional study using different topics in financial footnotes, could significantly add to the presented results.

Theme 2: A complementary study can add more inside to the artefact results by conducting the experiment by different experts in different countries.

Theme 3: Add more mapping capabilities to link some of textual information to their related figures in the main body of financial reports.

Theme 4: Applying not just textual disclosure of financial reports (footnotes) but other existing textual references extracted from internet, new, and etc. to the prototype to be able to increase efficiency and effectiveness of the results.
Fragen?

maryam.heidari@bwl.tu-freiberg.de