Exploring big data to sharpen financial sector risk assessment

David Roi Hardoon,
Monetary Authority of Singapore

---

1 This presentation was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.
EXPLORING BIG DATA TO SHARPEN FINANCIAL SECTOR RISK ASSESSMENT

“Building Pathways for Policy Making with Big Data”
BI-IFC / BIS International Seminar on Big Data, 26 July 2018

David R. Hardoon
Monetary Authority of Singapore
THE VISION

DOING BUSINESS AS UNUSUAL

https://www.denverpost.com
People remember….

- 80% of what they see and do
- 20% of what they read
- 10% of what they hear
## 3 Training Plans

### Interactive Data Analytics Course Catalogue

<table>
<thead>
<tr>
<th>Advanced</th>
<th>Intermediate</th>
<th>Basic</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Data Science Specialisation</td>
<td>- Intermediate Data Management &amp; Analysis</td>
<td>- Data Analytics: Core Concepts &amp; Techniques</td>
<td>- In-house</td>
</tr>
<tr>
<td>- Reproducible Data Science</td>
<td>- Data Preparation &amp; Exploration Using Python</td>
<td>- Data Analytics for Managers</td>
<td>- External - Online</td>
</tr>
<tr>
<td></td>
<td>- Intermediate Data Collection &amp; Analysis</td>
<td>- Customer Research through Surveys and Analytics</td>
<td>- Advanced Data Science: Machine Learning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Introduction to Data Management &amp; Analysis</td>
<td>- Building a Data Science Platform with Kubernetes, Anaconda &amp; JupyterHub</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Survey Questionnaire - Design &amp; Implementation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Introduction to Data Collection &amp; Analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Key Courses

- Introduction to Machine Learning
- Data Analyst - Visualisation II
- Intermediate Data Science (Python)
- Data Science for Managers: Crash Course in Data Science
- Starting a Data Science Project
- Customer Research through Surveys and Analytics
- Business Intelligence & Dashboarding
- Data Science & Artificial Intelligence for Senior Executives
- State Alive - A Foundation for Analytics
- Master Class on Building An Analytics Culture
- Basic Analytics for HR Professionals
- Driving Communications with a Data Analytics Strategy
CURRENT TECHNOLOGY & PROJECTS

MACHINE LEARNING

TEXT ANALYTICS & NATURAL LANGUAGE PROCESSING

VISUALISATION

AUTOMATION
Examples of Text Analytics and NLP

1. Scrape Chinese news websites
2. Clean data
3. Calculate sentiment scores
4. Visualise results and compute correlation
TEXT ANALYTICS & NATURAL LANGUAGE PROCESSING

Examples of Text Analytics and NLP

DATA

TOPIC MODELLING
(e.g. NMF, LDA)

TOPICS

<table>
<thead>
<tr>
<th>Topic 1</th>
<th>work processes, streamlining, prioritisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic 2</td>
<td>workload, headcount, resources</td>
</tr>
<tr>
<td>Topic 3</td>
<td>productivity, efficiency, systems and technology</td>
</tr>
</tbody>
</table>
VISUALISATION
Industry Wide Stress Test (IWST) Dashboard
Automating news surveillance

**PROBLEM**
Many departments in MAS currently filter news alerts by keywords, and rely on daily checks by support officers to assess if there are specific events that require attention. This process is time-consuming.

**AUTOMATION**
To help automate the process, we can train a model based on labelled articles.

**PRODUCTIVITY GAINS**
The model predicts whether a new article is relevant to a department. A similar model is currently used by Comms in filtering articles relevant to MAS, saving them a lot of time from the manual filtering.

**INPUT X**
News Articles

**OUTPUT Y**
- Relevant to Department
- Not relevant to Department
FUTURE OF MAS DIGITAL SUPERVISION

Example of ‘Bank A’ Portal View

- Latest news directly relevant to Bank A
- Ongoing communication btw MAS & Bank A
- Suspicious transactions reported by Bank A
- MAS’ assessment of Bank A’s risk over time
- Supervisor’s rating of Bank A on different indicators
- Inspection assessment of Bank A
ALERT PRIORITIZATION

The Pain & The Cure

130 Alerts per week
11.6% are important

4 Analysts
2 days per week

Machine to learn and classify alerts by importance

To capture as many important alerts as possible

75%

To reduce workload

25%
Workload Efficiency Gain
2 days per week reduced to 1.5 days. Monthly gain of 4 days

Alert Capture Recall
Objective is to capture as many positive alerts as possible

Note: The results are based on the optimal model and 10-fold cross validation
Enhance regulatory compliance with RegTech and latest technology

1. RegTechs to better understand the problem statements from FIs
2. FIs to better understand solutions from RegTechs
3. Encourage RegTech/FI interaction and potential POCs
ENABLING CROSS BORDER INSIGHTS
THANK YOU!

https://www.dreamstime.com