Predictive BI

Using the Past to Predict the Future

Antony Heljula
Technical Director
Using the Past to Predict the Future

• About Predictive BI
• The £1 Billion Problem
• How does it work?
• Oracle Products for Predictive BI
• Demonstration
• Summary
Using the Past to Predict the Future

About Predictive BI
About Predictive BI

What next?

- So you’ve built your dashboards...what do you do next?
  - Big Data?
  - Maps / Spatial?
  - Cloud?
  - More data sources?
  - Predictive BI?
About Predictive BI

Overview

• Predictive BI: *using the past to predict the future*

• Everyone can benefit from Predictive Analytics – it can add significant value to just about any department or organization

• All you need is:
  • A Business Problem
  • Historical data
About Predictive BI

“Obvious BI”

• How much Business Value do you get from a chart like this?
• It looks useful, but what are you going to do.... promote everyone?

Attrition Rate: Promotion vs No-Promotion
About Predictive BI
“Obvious BI”

• How much Business Value do you get from a chart like this?
  • It looks useful, but what are you going to do... promote everyone?

Key things I need to know:

Which departments are at most risk?
Who is the most likely to leave? Are they high-performers?
Are there other factors involved? e.g. salary, learning
What do I need to do in order to make them stay?
How much do we need to invest to reduce our risk?
About Predictive BI

**Example Scenarios**

• **Higher Education:**
  • Which students are high-risk of withdrawing early?
  • Which students are most likely to accept an offer?

• **Human Resources:**
  • Which employees are most likely to leave within a year?
  • Which employees are likely to be high-performers?
  • For a new job opening, which 10 employees are the most suitable?

• **Retail:**
  • Which products are most likely to become out of stock?
  • What is the likelihood a customer will churn?
  • What are the best-selling products going to be next weekend?
About Predictive BI

Where do you start?

• First of all, you need a Business Problem that can help you build up a solid Business Case...
Using the Past to Predict the Future

The £1 Billion Problem
The £1 Billion Problem

Some Background – HESA Published Statistics

- There are around 700K-800K full-time students entering Higher Education in England each year
  - Roughly 2 million students in total at any one time

- Course fees range from £7,500 - £9,000 per year

- Over 3 years, the total revenue per student can be between £23,500-£31,000 when you include accommodation fees

- For international students, the fees could be even higher
The £1 Billion Problem

Student Withdrawals – HESA Published Statistics

• Based on a sample of universities, 6.3% withdrew in their first year during 2010/11

• Do the maths!

• 700,000 x 6.3% = 44,100 students
• £23,500 average loss per student

Loss of funding = £1,036,350,000

An average of £7.5M for each of the 140 English Universities
The £1 Billion Problem

The HEFCE Classification

• One way to help reduce the number of early withdrawals is to understand which students are in a high-risk category

• Universities in the UK have a standard HEFCE Classification to give a Low/Medium/High risk rating for each student

• Let’s look at a real-life example....
The £1 Billion Problem
The HEFCE Classification – Real-Life Example

• Using real data from a UK University (1 faculty), here is the accuracy of the HEFCE Classification over 3 academic years (08/09 – 10/11)

  • Total number of students: 1259
  • Students with HEFCE Classification: 397 (35%)
  • Students with correct HEFCE Classification: 137
  • HEFCE Prediction Accuracy: 11%

Tossing a coin is 4x more accurate!

Only 1/3 students classified
The £1 Billion Problem

But it is not just Higher Education

• How much does it cost each time you lose an Employee?

• Average employee turnover rate in the UK is approx 15%

• An Oxford economics report (Feb 2014) estimates the average cost per employee is £30,614:
  • Lost “cost of lost output” whilst replacement employees get up to speed
  • The “logical lost” of recruiting and absorbing a new worker

• How many employees do you have?
  • 1,000 employees = £4.5M estimated cost of employee turnover
The £1 Billion Problem

How accurate are the predictions

• Example: A Higher Education organisation:
  • High/Low Risk students predicted to an accuracy of **83%**
  • (Previous method had an overall accuracy of just 11%)

• A global Financial Services company:
  • High/Low Risk employees predicted to an accuracy of **>90%**

• A global Consumer Goods company:
  • Out of stock products across 10,000 outlets predicted to an accuracy of **89%**
Using the Past to Predict the Future

How Does it Work?
How Does it Work?

Overview

1) Extract History
2) Build Predictive Model
3) Test Predictive Model
4) Use Predictive Model
5) Refine Predictive Model
How Does it Work?

1) Extract History

- Prepare an extract of history data, with as many attributes/metrics as possible (the more the merrier)
- Most importantly, provide a “Flag” to state the actual outcome

Here we have a “WITHDRAW_FLAG” which states for each Student whether or not they actually withdrew early from their courses.
How Does it Work?

2) Build Predictive Model

- We now pump the extracted data into the Data-Mining engine – it builds up a picture of what leads to “Y” or “N” scenarios.
How Does it Work?

3) Test Predictive Model

- You can now test your Predictive Model against another set of data, for example:
  - Use 80% of your data to build the model, the remain 20% is for testing it
  - Build your model using data for 2012/2013, test it against data for 2014

<table>
<thead>
<tr>
<th>Id</th>
<th>Prediction</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>Y</td>
<td>85%</td>
</tr>
<tr>
<td>124</td>
<td>Y</td>
<td>87%</td>
</tr>
<tr>
<td>125</td>
<td>N</td>
<td>93%</td>
</tr>
</tbody>
</table>

- Based on the results, you can choose to continue or go back to refine the model further until levels of accuracy are acceptable
How Does it Work?

4) Use Predictive Model

- You can now use the Predictive Model in your operational environments.

- For example:
  - Which employees are likely to leave in the next?
  - What happens if we increase their salary by 10%?

<table>
<thead>
<tr>
<th>Id</th>
<th>Prediction</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>123</td>
<td>Y</td>
<td>85%</td>
</tr>
<tr>
<td>124</td>
<td>Y</td>
<td>87%</td>
</tr>
<tr>
<td>125</td>
<td>N</td>
<td>93%</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
How Does it Work?

5) Refine Predictive Model

- At some point in the future you can refine your Predictive Model using the very latest data
- How often depends on the frequency of change in the source system
Using the Past to Predict the Future

Oracle Products for Predictive BI
Oracle Products for Predictive BI

Overview

• Oracle offers two exceptionally good products for delivering Predictive BI
  • Oracle Advanced Analytics
  • Oracle Real-Time Decisions (RTD)

• They are different architecturally – which one you need depends on your individual business need and setup

• Both have been proven to deliver accurate predictions
Oracle Products for Predictive BI
*Oracle Advanced Analytics*

- Historically, Oracle “Advanced Analytics” has been positioned for Predictive Analytics

- Oracle “Advanced Analytics” is formed of two different products:
  - Oracle Data-Mining
  - Oracle “R” Enterprise

- Embedded in the Oracle Database as an extra licensed option
Oracle Products for Predictive BI

**Oracle Advanced Analytics**

- **Oracle Data-Mining:**
  - More graphical, easier to set up and use
  - More automation and wizards available

- **Oracle “R” Enterprise Edition**
  - Requires scripts and coding logic
  - Historically used a lot with Research, Statistics and Higher Education organisations

- Licensed customers are free to choose which one to use – both deliver exceptional results
Oracle Products for Predictive BI

*Oracle Real-Time Decisions (RTD)*

- RTD is a middleware product, sitting outside the database
- It learns in real-time as each event takes place – you don’t have to refine your Predictive Models
Oracle Products for Predictive BI

Example

Session History

User Interactions

Sales Information

Recommendations

Data Warehouse
Oracle Products for Predictive BI

**Advanced Analytics vs RTD**

### Oracle Advanced Analytics

- More algorithms
- Data scientist has more scope for manipulating algorithms
- Needs a Data Scientist
- Requires Oracle Database EE
- Model refreshed periodically
- Better for “batch” learning
- Little integration OOTB
- More complex, longer to develop

### Oracle Real-Time Decisions

- No Data Scientist required
- Predictive Models built and maintained automatically
- Works with any data source
- Model learns in “real-time” – no need for refining models manually
- Better for frequently-changing models
- No history required to get started
- Self-Service integration capabilities
- Simpler, quicker to develop
Oracle Products for Predictive BI

Advanced Analytics vs RTD

Oracle Advanced Analytics

• More algorithms
• Data scientist has more scope for manipulating algorithms
• Needs a Data Scientist
• Requires Oracle Database EE
• Model refreshed periodically
• Better for “batch” learning
• Little integration OOTB
• More complex, longer to develop

Oracle Real-Time Decisions

• No Data Scientist required
• Predictive Models built and maintained automatically
• Works with any data source
• Model learns in “real-time” – no need for refining models manually
• Better for frequently-changing models
• No history required to get started
• Self-Service integration capabilities
• Simpler, quicker to develop
Oracle Products for Predictive BI

How much effort?

• Typically you would want to “try before you buy”: Perform a proof-of-concept on your own data to assess the value of Predictive BI

• A POC exercise normally takes around 4 weeks
  • 1 week data extract
  • 2 weeks building Predictive Models
  • 1 week analysis, presentation and next steps

• Can be delivered both on and off-site
  • Uses anonymised data – we do not need to know Ids, Names and Post Codes to do the profiling
Using the Past to Predict the Future

Demonstration
Using the Past to Predict the Future

Summary
Using the Past to Predict the Future

**Summary**

- The Business Value of Predictive BI is clear
  - Every student or employee saved equates to a definite amount of funding

- But there are other benefits, for example:
  - Improved satisfaction/engagement
  - Better reputation
  - Reduced churn to the benefit of competing companies / institutions
Helping Your Business Intelligence Journey