MARKET

ESPRIT Project Number EP 24456

IT Innovation

SPSS

Somerfield Stores

Public Final Report of ESPRIT HPCN PST Activity MARKET
Marketing Analysis for Retail using Knowledge Elicitation Tools

PAC/MARKET/D0.13 Version 2

Paul Allen, IT Innovation
Trevor Lilley, SPSS
Jerry Warren, Somerfield Stores

11 April 2000

Abstract

The MARKET project demonstrates how a state-of-the-art data mining (knowledge discovery) toolset combined with a large data warehouse can be effectively deployed into a retail user environment which is not staffed by IT specialists.

Synopsis

This report is the final report of the MARKET Project (ESPRIT EP 24456). The project involves the following partners: IT Innovation (UK), SPSS Inc (UK) and Somerfield Stores Ltd (UK). The project was co-ordinated by IT Innovation and further information may be obtained from:

Paul Allen
Tel: +44 23 80 76 08 34
Fax: +44 23 80 76 08 33
mailto: pja@it-innovation.soton.ac.uk

Data mining is a maturing technology and application area. However, there are few public case studies demonstrating the power and benefit of the technique. It is often viewed as an arcane 'black art' requiring highly skilled IT specialists to exploit it. In the MARKET project we set out to demonstrate how non-IT specialists using data mining techniques can derive business benefit from a large corporate data warehouse.

Somerfield have investigated their bread ordering process. The existing process is based on demand, which is well understood under normal conditions, but not during promotional periods. Bread is particularly significant for the following reasons.

- Bread has a short shelf life. Over-ordering leads to increased wastage and loss of revenue.
- Bread is a staple product and is contained in 75% of baskets. Under-ordering leads to loss of sales, customer satisfaction and market share.

Two main effects that cause the demand of bread to be complex during promotional periods are domino and steal. The concepts were already intuitively understood.

- ***Steal*** is when sales of a particular product are affected due to a promotion on another similar product.
- ***Domino*** is when customers switch to an alternative product if their first choice product is unavailable. This may cause the alternative product to sell out, shifting demand to a third product, and so on.

However these effects were not quantitatively understood and were therefore not able to be incorporated into the bread ordering process.

In the MARKET project we have shown how neural network technology can be used for basic demand forecasting. We have shown how association rules and other data mining techniques can be used to understand how customers purchase bread and how complex effects such as steal and domino inter-operate.

This has resulted in an improved bread ordering algorithm for Somerfield and has allowed them to implement a replicable process for data mining. Somerfield are now able to implement processes to enable more accurate bread stock levels, resulting in fewer dissatisfied customers and reduced wastage. They have an improved understanding of their business and an enhanced business intelligence capability that they have already exploited in other areas of their business.

Executive Summary

Data mining is a maturing technology and application area. However, there are few public case studies demonstrating the power and benefit of the technique. It is often viewed as an arcane 'black art' requiring highly skilled IT specialists to exploit it. In the MARKET project we set out to demonstrate how non-IT specialists using data mining techniques can derive business benefit from a large corporate data warehouse.

The objectives of the MARKET project were to:

- demonstrate the potential for adding value to the information assets of retail organisations through effective knowledge discovery from large data warehouses;
- demonstrate that end-users in a retail marketing department can use a state-of-the-art knowledge discovery toolset to effectively and productively generate actionable high-value information from a large data warehouse on a high-performance platform; and
- generate a body of large-scale knowledge discovery experience, from the retail sector, which will form the basis for generic dissemination and specific marketing activities.

In order to achieve the objectives of the MARKET project we have undertaken a major data mining exercise with Somerfield, using a recently published standard cross-industry process (CRISP-DM).

Somerfield have explored the issues associated with the adoption of data mining as a business intelligence capability and have evaluated this through independent data mining activities which have been undertaken elsewhere in the business.

We have built a proof-of-concept high-performance validation facility. This approaches uses the standard Clementine method to explore the data and derive hypothesis, but allows the user to validate his hypothesis using the power of the warehouse.

In order to increase take-up we have constructed a Web-based public demonstrator based on the Somerfield data mining activities.

Somerfield have investigated their bread ordering process. The existing process is based on demand, which is well understood under normal conditions, but not during promotional periods.

Two main effects that cause the demand of bread to be complex during promotional periods are domino and steal. The concepts were already intuitively understood. However these effects were not quantitatively understood and were therefore not able to be incorporated into the bread ordering process.

In the MARKET project we have shown how neural network technology can be used for basic demand forecasting. We have shown how association rules and other data mining techniques can be used to understand how customers purchase bread and how complex effects such as steal and domino inter-operate.

This has resulted in an improved bread ordering algorithm for Somerfield and has allowed them to implement a replicable process for data mining. Somerfield are now able to implement processes to enable more accurate bread stock levels, resulting in fewer dissatisfied customers and reduced wastage. They have an improved understanding of their business and an enhanced business intelligence capability that they have already exploited in other areas of the business.

Contents

1 State of the Art

1.1 Information management for food retailing

Over recent years the food retailing sector throughout Europe has become highly competitive. Advanced IT has become an increasingly important factor in the creation of competitive advantage for companies in this sector, with substantial investments being made across the industry in in-store Electronic Point Of Sale (EPOS) systems, electronic commerce systems and in sophisticated management information systems. However, the adoption and exploitation of advanced IT is not universally recognised in what is a relatively conservative sector and awareness raising, case studies and best practice information are required to increase take-up.

1.2 Data warehousing

During the years preceding the MARKET project there has been a rapid growth of interest in data warehousing, largely due to the economics and availability of technology able to provide both high levels of processing power and huge storage capacities. The availability of client-server technology and parallel HPC platforms capable of running parallel versions of relational database management system (RDBMS) software, has made data warehousing an increasingly affordable, and attractive, option to many companies.

At the start of the project many organisations had built, or were building data warehouses. However, the majority of the analysis carried out against data warehouses was simple querying and OLAP-style reporting. When the decision maker is looking for patterns with *a priori* unknown characteristics, standard query generation techniques, OLAP (On-Line Analytical Processing) or graphical tools are not applicable. Few organisations had explored the more powerful analysis offered by data mining and knowledge discovery techniques.

1.3 Basket level sales analysis

Many retailers have historically provided their management with sales information aggregated at the store level.

A recent major industry trend has been to enable the analysis of sales data at the basket level, i.e. to retain information on every item purchased and to know which items were purchased on the same transaction. The data volumes associated with basket sales analysis run to 100s of Gbytes for a retailer operating on the scale of Somerfield. It is only with the current generation of IT that it is becoming viable to undertake such an analysis. The extra information contained within basket data enable much richer analysis to be performed, particularly using data mining techniques.

1.4 Knowledge discovery and data mining

At the start of the project there was much hype about the data warehousing marketplace as a whole, and about data mining in particular. Much work had been performed by academic and industrial research groups to refine the basic techniques of data mining, and a number of specialist vendors offered data mining tools. However, there existed no replicable process for user organisations to implement data mining in a structured way and few case studies to justify investment in the approach.

Many data mining tools were client based, which extract and maintain local copies of data. These tools enable powerful data mining to be performed on subsets of the data. However, they do not harness the full high performance computing and networking (HPCN) power typically available in data warehouse platforms.

1.5 Somerfield Stores information systems

Somerfield is a major UK food retailer which, at the start of the MARKET project, had about 625 stores, selling over 12,000 product lines, employing 48,000 staff, serving an average of 7.7 million customers per week, and having an annual turnover of £3.2 billion.

The data warehouse was implemented during 1996 and contained the following information: sales; wastage; issues (stock movements to and between stores); baskets; and reference data.

Somerfield did not have any experience of data mining. Most analysis was performed with spreadsheets and simple reporting packages.

1.6 SPSS Clementine

SPSS's Clementine knowledge discovery product provides a toolkit of techniques within a client-based environment designed to enable the non-IT specialist to effectively discover hidden knowledge.

1.7 Objectives

The objectives of the MARKET project were to:

- demonstrate the potential for adding value to the information assets of retail organisations through effective knowledge discovery from large data warehouses;
- demonstrate that end-users in a retail marketing department can use a state-of-the-art knowledge discovery toolset to effectively and productively generate actionable high-value information from a large data warehouse on a high-performance platform; and
- generate a body of large-scale knowledge discovery experience, from the retail sector, which will form the basis for generic dissemination and specific marketing activities.

1.8 Partners

The MARKET project involved the following partners.

- **Somerfield Stores Ltd (UK)** are a major UK supermarket chain and have provided the end user role.
- **SPSS Inc (formerly ISL Ltd) (UK)** produce the data mining package "Clementine" and have provided the data mining tool and associated consultancy.
- **IT Innovation (formerly PAC) (UK)** is a systems integration organisation specialising in the innovative application of information technology, and has played a software development, technical consultancy and management role in the project.

2 Approach taken

In order to achieve the objectives of the MARKET project we have undertaken a major data mining exercise with Somerfield, using a recently published standard cross-industry process (CRISP-DM). This exercise has been used as the basis of a public demonstrator which is available on the Web.

2.1 Data mining

2.1.1 Actual projects undertaken

An initial 'trial' data mining exercise investigating the effects of putting goods as "reduced to clear" was undertaken. This was done in order to provide an initial and early validation of the approach and in order to generate awareness and understanding within Somerfield.

Within the supply chain department, Somerfield have investigated the bread ordering process in order to improve bread availability and reduce wastage. This investigation has formed the major effort within the MARKET project.

2.2 Business intelligence capability

Somerfield have investigated the wider issues of adopting data mining as a business intelligence capability.

Within the marketing department, Somerfield have investigated home delivery and home shopping issues. This investigation has been *enabled* by the MARKET project and has been carried out partly within, and partly subsequent to, the project.

2.2.1 Lifecycle

In the MARKET project, a draft version of the CRISP-DM data mining process has been used. CRISP-DM is a generic cross-industry process for data mining that has been created in the EC project CRISP-DM. In each data mining exercise, the CRISP-DM processes have been addressed in the following way:

- **Business Understanding:** *This initial phase focuses on understanding the project objectives and requirements from a business perspective, and then converting this knowledge into a data mining problem definition, and a preliminary plan designed to achieve the objectives.*

  This phase was achieved through issuing a briefing document to business users, followed up by brainstorming meetings involving IT and data mining consultancy personnel. The business analyst drew up a preliminary plan.

- **Data Understanding:** *The data understanding phase starts with an initial data collection and proceeds with activities in order to get familiar with the data, to identify data quality problems, to discover first insights into the data, or to detect interesting subsets to form hypotheses for hidden information.*

  This phase was achieved using the flexible graphical interface provided by Clementine to explore the data.

- **Data Preparation:** *The data preparation phase covers all activities to construct the final data set from the initial raw data.*

Data preparation was achieved using end user client tools to extract data from the warehouse. Clementine was used to transform the data using the various transformation operations provided.

- **Modelling:** *In this phase, various modelling techniques are selected and applied, and their parameters are calibrated to optimal values.*

  The Clementine package contains a variety of modelling algorithms such as neural networks, rule induction and association rule algorithms. The use of these algorithms was explored in the MARKET project.

- **Evaluation:** *In this phase the model is more thoroughly evaluated and the steps used to construct the model are reviewed, to be certain it properly achieves the business objectives.*

  The results of the data mining were presented to, and reviewed by, senior business personnel. They could then highlight issues for further investigation which, when resolved, enable buy-in to the results at the highest level.

- **Deployment:** *This phase concentrates on making use of the knowledge gained, and can be as simple as generating a report or as complex as implementing a repeatable data mining process*.

  Once the data mining results were presented as a standalone report, Somerfield have been able to deploy the results using standard mechanisms. Somerfield are currently implementing a repeatable data mining process.

2.2.2 End-users

A business analyst from the supply chain department carried out the bread availability analysis. The business analyst participated in the standard basic and advanced Clementine courses and worked with SPSS staff during the period of analysis. The reason for this interactive involvement was to transfer a great deal of expertise and to set up the analyst as an internal expert within Somerfield.

A business analyst from the marketing department, who has attended Clementine courses but has not worked interactively with SPSS staff during analysis carried out the home delivery and home shopping analysis.

2.2.3 Success criteria and evaluation

During the requirements capture phase, the following success criteria were identified.

- Positive and conclusive results should be obtained from the present data, from the use of Clementine
- The interpretation of Clementine analysis should be commercially sound and provide readable data and graphics.
- The data mining tool, Clementine, should be user friendly to, and usable by, non-expert staff with sufficient training.
- The tool and the data mining process should be supportable by the business.
- The principle of data mining should be accepted by the business.

2.3 Demonstrator

The approach taken was to build a Web demonstrator, based around the analysis performed in the project, to highlight the issues that were investigated.

The demonstrator was based on data provided by Somerfield, transformed and augmented as appropriate to result in a demonstrator for a fictitious supermarket. In this way, a realistic demonstrator can be produced that does not release business-sensitive information. The scope of the demonstrator covers the business problem, the data understanding, data preparation, modelling and deployment stages of the data mining process. The demonstrator is designed to be viewed by a business analyst or executive in approximately fifteen minutes.

3 Results, Achievements and Benefits

3.1 Bread

3.1.1 Success Story

Somerfield have investigated their bread ordering process. The existing process is based on demand, which is well understood under normal conditions, but not during promotional periods. Bread is particularly significant for the following reasons.

- Bread has a short shelf life. Over-ordering leads to increased wastage and loss of revenue.
- Bread is a staple product and is contained in 75% of baskets. Under-ordering leads to loss of sales, customer satisfaction and market share.

Two main effects that cause the demand of bread to be complex during promotional periods are domino and steal. The concepts were already intuitively understood.

- ***Steal*** is when sales of a particular product are affected due to a promotion on another similar product.
- ***Domino*** is when customers switch to an alternative product if their first choice product is unavailable. This may cause the alternative product to sell out, shifting demand to a third product, and so on.

However these effects were not quantitatively understood and therefore not usable by the bread ordering process.

In the MARKET project we have shown how neural network technology can be used for basic demand forecasting. We have shown how association rules and other data mining techniques can be used to understand how customers purchase bread and how complex effects such as steal and domino inter-operate.

Business benefits

The data mining activity in MARKET has resulted in an improved bread ordering algorithm for Somerfield. Somerfield are now able to implement processes to enable more accurate bread stock levels, resulting in fewer dissatisfied customers and reduced wastage. They have an improved understanding of their business and an enhanced business intelligence capability that they have already exploited in other areas of the business.

3.1.2 Discussion

The data mining exercise has led Somerfield to a greater understanding of effects that they were not previously aware of. They now have a much better picture of how different types of customer purchase bread.

The use of neural networks for forecasting has been a useful result. Neural networks have proven slightly more accurate than existing methods for forecasting demand during 'standard' periods - nevertheless, they have also verified the basic reliability of the existing ordering methods.

Figure 1 shows the output of the neural network. This compares the actual sales (blue) with the forecast sales (green), together with the rolling average (red). The error is shown in purple.

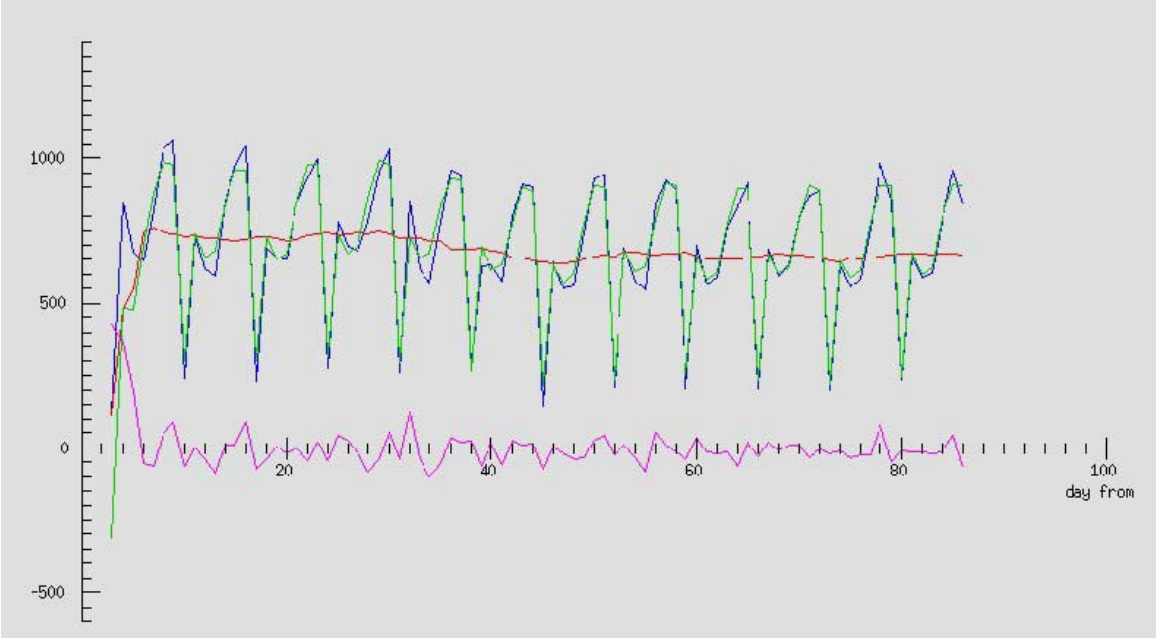

Figure 1 Neural network modelling for demand forecasting

The main benefit has been through the investigation of domino and steal effects. Before the MARKET project these effects were known to happen in general terms, but the interactions between specific products were not known. The use of data mining has shown which products steal from which other products and by how much. It has shown specifically how other products are affected when certain products sell out. The greater understanding of domino and steal effects have resulted in the creation of rules that can be used to enhance the bread ordering process.

The interpretation of the bread availability analysis is commercially sound and has been identified at a senior level as being extremely important for Somerfield. A trial has been planned involving 12 stores. There will be 6 stores which will use the new bread ordering algorithm and 6 stores which will act as a control during the course of the experiment. At the end of the trial, Somerfield will be able to calculate the effect of the new algorithm in terms of reduced wastage. They will also investigate the effects on sales and availability.

3.2 Marketing

Data mining has been used in the justification for, and development of, Somerfield's home delivery and home shopping programme.

At any early stage, data mining was used to analyse initial home-delivery trial data to conclude that in certain stores the home delivery programme resulted in an increase in market share. Rule induction was used to derive rules to describe the type of stores in which home delivery would be most beneficial to customers. These stores turned out to be orthogonal to Somerfield's existing store classification and were therefore not derivable from their standard methods.

In the development of Somerfield's home shopping programme, data mining has produced positive results. Current global best practice indicates that it is not desirable to sell the entire product range through home shopping channels, and that it is desirable to select a reduced range. Data mining techniques have been used to assist in the selection

of the most favourable products for home shopping. Additionally, associations have been used to explore the potential of linked offers through home shopping.

The "24 by 7" home shopping programme is currently live via digital TV and has been announced on the Web at http://www.24-7.co.uk/. This 24 hour depot-based home shopping service will establish Somerfield as the leading home shopping supermarket in the UK.

Business benefits

The home delivery and home shopping analysis has already enabled Somerfield to obtain new and industrially important insights, and has been a factor in Somerfield's strategic push in this direction, as described in their press statement of November 1999 (http://www.somerfield.co.uk/more/presslatest.asp?release=11Nov99.html). The data mining activities not only formed part of the justification for the launching of these major e-commerce initiatives, but also indicated how they should be implemented. This will result in significantly increased competitiveness for Somerfield.

3.3 Business Intelligence Capability

One important finding has been that it has been possible, using Clementine, to verify knowledge that had previously only been assumed. Using intuition and business knowledge, Somerfield business analysts have been able to formulate rules and demonstrate effects graphically. In particular it has become possible to investigate to a finer level of granularity than previously possible. Data mining has verified, in a fraction of the time, results that have previously been calculated using spreadsheet analysis.

Clementine is user-friendly and powerful and is suitable for deployment as a business intelligence capability. It is capable of fast 'broad' analysis, with subsequent drill down to a finer level of detail.

The success of the marketing analysis using Clementine demonstrates the usability of the tool, and the usability of the data mining process, by non-expert staff at Somerfield.

Somerfield are implementing support for data mining within the data warehouse capability. This includes support for HPC mining using three-tier technology and internal training mechanisms involving power users providing first line support for conventional users.

Business benefits

Somerfield's enhanced business intelligence capability means that they can reap the benefits from MARKET in new ways, after the project has finished.

3.4 HPC validation

In the MARKET project we have built a proof-of-concept high-performance validation facility. This uses the standard Clementine method to explore the data and derive hypothesis, but allows the user to validate his hypothesis against the full data volumes on the data warehouse.

We have built a validation module that interfaces to Clementine and generates optimised Oracle-based SQL to validate C5.0 decision trees. This approach was selected because of the widespread use of Oracle on data warehousing platforms (including Somerfield's) and the suitability and widespread take-up of the C5.0 algorithm.

In parallel with this activity, SPSS have been developing their CHESS server (Clementine High Efficiency Scalable Server) which provides a high-performance middle tier to exploit the power in the warehouse for data preparation (but not validation).

The high performance validation capability will allow the business analyst to verify their hypothesis on the entire data set. This will give them more confidence in their assertions and facilitate the business exploitation of the analysis.

3.5 Demonstrator

The demonstrator has been implemented as a set of interactive Web HTML pages, based around the bread availability analysis. It demonstrates a fictitious supermarket that is investigating their bread ordering process and contains the following stages:

- data understanding stream and results;
- initial modelling of non-promotional periods using neural networks;
- investigation of domino effects;
- investigation of steal effects; and
- impact on bread ordering process.

Figure 2 shows a Web page from the demonstrator. The demonstrator can be found at http://www.hpcn-ttn.org/.

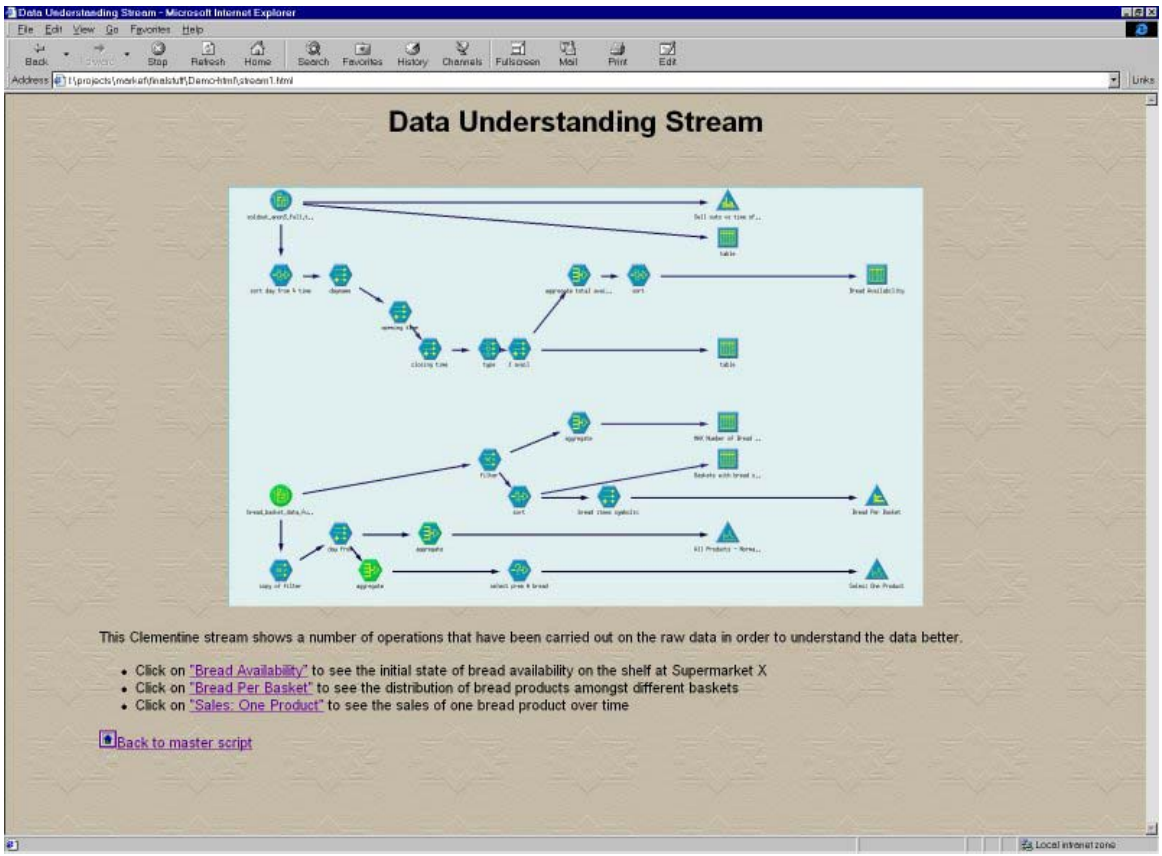

Figure 2 A demonstrator Web page

3.6 Comparison with objectives

The MARKET project has:

- demonstrated the potential for adding value to the information assets of retail organisations - this is illustrated by the bread and marketing success stories;
- demonstrated that end-users in a retail marketing department case use state-of-the-art knowledge toolsets - this is illustrated by Somerfield's enhanced business intelligence capability; and
- generated a body of large-scale knowledge discovery experience - this is illustrated by the public demonstrator and dissemination material.

3.7 Lessons learned

An important factor is the characteristics of the end-user who is performing the data mining. The data-mining user should have the following qualities.

- They should have a good knowledge of the business in which they operate. This is essential in order to find significant and commercially sound results.
- They should have an analytical perspective and be able to generate and follow-through hypotheses.
- They should be IT-literate – it is not necessary to have an IT expert, but the user should be comfortable with desktop technology.

An ideal user is an IT-oriented business user. Alternatively business-oriented IT users can also be effective. Data mining can be a complex task and some degree of immersion on the part of the analyst user is required. Ideally this would involve some periods of full-time activity for a number of weeks, although half-time is sufficient.

It is useful to train an expert within the user organisation through participation in organised courses and through interactive working with an external data mining expert. Contact with external expertise is particularly beneficial:

- to empower the user organisation with a data mining capability; and
- to provide a validation and troubleshooting role.

The MARKET project has highlighted the fact the quantity and quality of the raw data is key to the data mining process. Subsets of the data have been extracted using an end-user analysis tool. This proved to be possible, although inconvenient and time-consuming, due to the fact that the tool was not designed to be used for high volume data extractions. The requirement for much larger data volumes in order to validate the analysis has been identified. This process is much more efficient using HPC-enabled technology such as SPSS CHESS.

4 Dissemination

The dissemination in the MARKET project was aimed at business analysts, business executives and IT strategists in the retail sector. This section lists the major dissemination achievements.

4.1 Publications and the WWW

- "Clementine is in the MARKET for European retail solutions"
  CLUG News
  April 1998
  Page 7
- "MARKET: Somerfield's use of data mining"
  PriceWaterhouseCoopers web site
  June 1998
- "Bread Dominoes"
  Stratagem
  July 1998
- "Somerfield Stores Generates New Insights with Clementine"
  DM Review Magazine
  July 1998
- "Key players and their products: let's go mining"
  Knowledge Management
  A Learned Information Publication
  December 1998/January 1999
  Page 32

4.2 Conferences

- "Data Mining: A Retail Case Study"
  Flyer distributed from IT Innovation exhibition stand
  Business Intelligence 98
  London
  June 1998
- "Applications of Data Mining Techniques in Marketing, Retail, Finance and Insurance Industries"
  Presented by Colin Shearer, SPSS
  Unicom Seminar
  Olympia, London
  9-10 December 1998
- "Data Mining – A Retail Case Study"
  Presented by Stephen Hellberg, IT Innovation
  HPCN TTN Belfast Workshop on Data Mining
  Queen's College, Belfast
  2 March 1999
- "Retail Applications of Data Mining"
  Presented by Simon Metcalfe, Somerfield
  Business Intelligence 99

Olympia II, London
16 June 1999

4.3 Demonstrator

The public demonstrator can be viewed at http://www.hpcn-ttn.org/

5 Conclusions

We have shown how neural network technology can be used for basic demand forecasting. We have shown how association rules and other data mining techniques can be used to understand how customers purchase bread and how complex effects such as steal and domino inter-operate.

The data mining activity in MARKET has resulted in an improved bread ordering algorithm for Somerfield. Somerfield are now able to implement processes to enable more accurate bread stock levels, resulting in fewer dissatisfied customers and reduced wastage. They have an improved understanding of their business and an enhanced business intelligence capability that they have already exploited in other areas of the business.

Data mining has been used in the justification for, and development of, Somerfield's home delivery and home shopping programme. They have concluded that in certain stores the home delivery programme resulted in an increase in market share. Data mining techniques have been used to assist in the selection of the most favourable products for home shopping and to explore the potential of linked offers through home shopping.

The home delivery and home shopping analysis has already enabled Somerfield to obtain new and industrially important insights and has been a factor in Somerfield's strategic push in this direction. The data mining activities not only formed part of the justification for the launching of these major e-commerce initiatives, but also indicated how they should be implemented. This will result in significantly increased competitiveness for Somerfield.

Somerfield have verified all five of the success criteria identified in the project. In particular:

- positive and conclusive results have been obtained from the data, as illustrated by the success stories;
- the interpretation of the analysis has been shown to be commercially sound, and has been a contributing factor in the launch of a major new e-commerce activity;
- the data-mining tool is usable by non-expert staff with sufficient training – the marketing analysis was performed entirely by end-users from the marketing department;
- the tool and the data mining process is supportable by the business and is being integrated with Somerfield's data warehouse strategy; and
- the principle of data mining is accepted by the business and is now an accepted method for creating and justifying new opportunities.

5.1 Exploitation

SPSS now have an improved understanding of the retail sector and of the issues associated with basket analysis and supply chain forecasting. The MARKET project will be an excellent case study to enable SPSS to sell high performance data mining products and services to the retail sector.

IT Innovation now have an improved understanding of the retail sector and of the issues associated with basket analysis and supply chain forecasting. The MARKET

project will be an excellent case study to enable IT Innovation to sell innovative data mining and knowledge management services to the retail sector.

Somerfield are implementing support for data mining within the data warehouse capability. This includes support for HPC mining using three-tier technology and internal training mechanisms involving power users providing first line support for conventional users. Somerfield intend to realise benefits of data mining in many different areas throughout the organisation.

6. Contact Details

Company	Contact	URL	Role
IT Innovation Centre	Paul Allen	http://www.it-innovation.soton.ac.uk/	Project management
SPSS Inc	Trevor Lilley	http://www.spss.com/clementine/	Data mining expertise
Somerfield Stores Ltd	Jerry Warren	http://www.somerfield.co.uk	End-user

The project ran for 22 months, starting in September 1997 and ending in June 1999. The European Commission contributed 230,500 Euro to the project.