Managerial Support Systems

Alex Sverdlov
alex@theparticle.com

1 Introduction

Besides the systems that support operational needs of businesses (such as enterprise resource planning, customer relationship management, etc.), there are a wide variety of systems that are more specialized to serving the needs of management. These can very broadly be categorized under *decision support systems*. These range in complexity from a simple Excel spreadsheet, to complex models that are trained on historical data.

Such decision support systems often consume operational data from enterprise systems, but the purpose is to summarize and to build models, etc., not to run the everyday business. Their primary purpose is to assist management with making decisions that improve business.

2 Data

The traditional view is that businesses measure and record everything that is required to run the business (or required by regulation). Modern businesses often take a different approach: measure and record everything that is measurable, useful or not. What is ultimately useful is not known at the time of data collection, and with data storage being so cheap, it makes business sense to just record everything.

Data has value. More data has more value. Even if something is not immediately useful, someone may find a use for it later. Or perhaps the data may be sold for revenue, etc.

This perspective naturally leads to data overload. Everything that is going on within the company or is observable by the company ends up in the log.

3 Decision Support Systems

Management needs an actionable summary view into all the data that is collected. Often these are reports or dashboards that display key business attributes, major positions, risks, etc.

On a small scale, such tasks may be accomplished by Excel. Management can cut & paste previous quarter numbers into Excel, apply regression or some other projection and get a prediction of what next quarter numbers may look like. They can also play what-if games by adjusting cell values with educated guesses to get (perhaps actionable) results.

Once things get beyond a certain point, it is not practical to aggregate such data in a desktop application, much less to run models in that environment.

Within bigger enterprises, reports and dashboards are created to present key business metrics in a way that enables management to quickly grasp what is happening. Each report or dashboard may highlight a particular perspective of the business—and may be built and consumed independently of other reports.

4 Proxy Features

Management reports may include things that are not observable (measurable). In such scenarios, a proxy measurable may be used in place of a nebulous concepts such as "customer satisfaction", or "user engagement".

For example, the count of purchases and the amount of money spent could be a good proxy measure for customer satisfaction, and number of replies or new posts may be a good proxy measure for user engagement.

5 Data Mining & Data Science & AI

Decisions must be backed by data. If there is no data to make a decision, then the first priority should be to gather data. Making decisions without data is stupid.

Data mining is the process of extracting information out of data. This can take on multiple meanings, from simple statistics to building prediction models.

Summary statistics, such as totals, averages, standard deviations, medians, interquartile range (IQR) are often the first steps to summarizing data.

Recently a fancy title has emerged: *Data Science*. This is essentially applying the scientific method on enterprise data: kind of formalized data mining. The key thing with data science is building models using data, and then running experiments trying to find places where the models break. Models that withstand attacks (are confirmed by data) gain more confidence. Those models are then used to make predictions or model some aspect of the enterprise.

5.1 Labeled Data

There is a huge problem in data science: the lack of labeled data. Very often, labels are assigned by experts, analysts, etc., and that is an expensive and error-prone process, that yields comparatively small amounts of labeled records.

Most data that businesses deal with is not labeled.

One way to deal with this issue is to let time label the records. If we need a model to assign credit risk to customers, we can use historical data—what the customer provided in their credit application, and how did that customer perform in a year after that application.

This has to be done very carefully: as customers who did not get credit are not part of this model. One approach is to *randomly* grant customers credit, and then use only those customers to build a model. But randomly giving out loans is obviously an expensive and risky proposition.

5.2 Distance Measure

Another major problem in all such systems is how to measure distance. Many similarity measures (either between customers, suppliers, risks, etc.,) relies on a mathematical distance measure—such as Cartesian distance. The problem is that the mathematical distance measure may be meaningless in regards to the business.

For example, what does it mean when distance between customer A and customer B is 7? What dimensions are being compared? Is it customer age? Customer address? Customer purchase history? Customer gender? Customer return counts, etc.? A single number is not detailed enough to capture that detail—and yet this is exactly the metric that is used by a lot of models that would cluster or label customers.

5.3 Machine Learning

Machine learning is the technical guts of what data scientists do. Essentially we need model of the business—and since we are dealing with data, we need a model of the data. A model is a concise general description of the data.

For example, the data may be a thousand observations, and our model (description) could just a hyperplane. This does not mean that the business or data are a hyperplane—just that this description is enough for some purpose.

Machine learning is mostly interested in various models that can be trained on data using various techniques. Most of the models are "simple"—lines, planes, spheres, etc., and various collections of the above. Being simple, they are often easy to train from examples, and fast to execute.

5.4 Artificial Intelligence

The holy grail of general purpose AI is still very far away, but some applications of machine learning have made great strides in limited well defined tasks.

These include: speech recognition, object recognition, OCR.

Areas that have not experienced growth are generalization: re-purposing a model from one task to another is still very illusive—something humans do very well.

Many resource allocation tasks are *hard*, and it is not clear if AI will make them any easier (e.g. deciding where to allocate capital).

5.5 Expert Systems

Expert systems are automated expert rules of thumb. Suppose we need to diagnose a problem. One way is to ask an expert about what they would examine, and what conclusions they would draw. This rule set may be codified in *if-then-else* format, such that the system can walk through these rules diagnose the problem in the same way that an expert might.

There is a limitation: experts are expensive, and getting them to exhaustively list every conceivable situation is difficult.

Expert systems have been shown to work (not necessarily work well) in well defined limited domains—if we can proceed with a few well placed questions/answers, then expert systems might work as well as an expert.

5.6 Decision Trees

Decision trees are kind of an automated way of building an expert system. By examining labeled data, the computer comes up with the best questions to ask in order to arrive at the target label. Every question/answer in a sequence maximizes the information gain measure.

Straight decision trees tend to over-fit the data, meaning as performance on training set increases, the performance on the test set decreases (the model is memorizing the training data).

One solution to this is to stop training early.

Another solution is to over-train the model, and then as a subsequent iteration, remove any rules that do not degrade performance.

Yet another solution is to train many shallow trees, with the result being the majority vote of such simpler models.

5.7 Neural Networks

Neural Networks are often seen as a model of how the brain operates.

In recent years, with modern architectures (deep-learning, convolution neural networks, ReLU activation function, back-propagation, etc.), neural networks have managed to perform tasks that were previously seen as very difficult. Such as identifying objects in images and video, reliably recognizing speech, etc.

While they appear to be able to solve most problems folks have thrown at them, they do not generalize well—and it is not exactly clear in what scenarios the results would break down.

The other big downside is that they require a lot of training data—while human beings often learn from one or two instances, neural networks often require thousands.

6 Knowledge Management

Another critical component of management support system is knowledge management. Every enterprise generates knowledge. Either created by individuals or teams, etc. When an employee becomes an expert at what they do, they have gained knowledge that is part of the company assets—the worker is more valuable than an identical worker without the knowledge.

While some of this knowledge may not be transferable, or easily written down, there are many other kinds of knowledge that the company should make an effort to organize and retain. The expertise gained should not reside in the heads of individuals, but documented and shared such that future teams can benefit.

Wiki pages are one way to share organization level knowledge: employees can creates pages that summarize the problem and the solutions they have tried, and more importantly, which solutions worked.