

DATAMINE 軟件時，又需要輸入一些數據，才能啓動工序，而所得到的結果又沒有直接導致這個原因多數是用家不明白每個工序和參數的用途。其實輸入參數有助引導工在有限的時間裏尋找有用的資料。自動化的 MINING 在於你能否了解工序，配合發揮 DATAMINE 的能力，從而得到真正



DATAMINE 至今讀者看到一些字眼，例如分類、拼法條等。它們都是在 MINING 技術中出來的，可以技術的應用是非常普遍，一半 DATAMINE 軟件開發公司，技術加入其系統中，這技術就樹 (DECISION TREE)。由應用廣泛，所以清楚掌握決策作、優點和缺點，不但可以為擇合用的軟件，而且在應用這時候，能可避重就輕，大大發 DATAMINE 的功力。

讀決策樹，顧名思義就是將整的流程，把其中可能牽涉到的件，有系統地以樹狀的結構呈。決策樹是由樹根、樹節和樹成 (見圖一)。樹根是分析的樹葉是終點，樹節是分差路 (決策點)。

### 馬作例設定決策樹

現在用一個數據庫共三百三十二，資料取材於香港賽馬紀錄，用其中一場參賽馬 (共十四過往賽蹟 (見表一)。現在來決策樹的運作。首先決策樹演式放在起點的樹根，當用家確定分析的特徵後，演式把其餘的特

一輪的分裂中被選用組過有能力的特徵環境，直到樹葉出現，此時的特徵會形成。

數值種類可分為文字和數字。字種類比數字種類簡單，特徵值是個個體，沒有先後之分，因此可以分開或組合。文字種類可以分開或組合，文字種類要分為「不連續型」和「連續型」。馬匹的馬匹位特徵是「不連續型」的數字種類，而連續型的特徵值需要找一個分割點，統計學的標準差可以作為分割基礎。分割後的每一小組，所得的標準差是最小。決定數字種類是非常重要的，若果把馬匹位特徵作為分析對象，連續型的數字所得到的答案是：少於五數字為一組，大過或等於五數字為另一組。而不連續型的數字所得到的答案是：一、二、四、十一、十三、十四為一組，三、五、六、九為另一組，七、八、十為最後一組。重要的是分析結果的可用性，筆者覺得連續型的數字種類比較適合馬匹位特徵。

如何顯示分析目錄  
變成選擇數據庫，可以開始去

# Data Mining Versus Semantic Web



Veljko Milutinovic, [vm@etf.bg.ac.yu](mailto:vm@etf.bg.ac.yu)

<http://galeb.etf.bg.ac.yu/vm>

This material was developed with financial help of the WUSA fund of Austria.

# DataMining versus SemanticWeb



- Two different avenues leading to the same goal!
- The goal:  
Efficient retrieval of knowledge,  
from large compact or distributed databases,  
or the Internet
- What is the knowledge:  
Synergistic interaction of information (data)  
and its relationships (correlations).
- The major difference:  
Placement of complexity!



DATAMINE 軟件時，又會導致  
能啓動工序，而所得到的結果又沒有  
導致這個原因多數是用家不明白每個工序  
和參數的用途。其實輸入參數有助引導工  
有限的時間裏尋找有用的資料。自動化的  
MINING 至於你能否了解工  
彈 DATAMING 的能力，從  
的實  
DATAMING 至今讀者  
看到一些字眼，例如分類、排  
法條等。  
MINING 技術中出  
技術的應用是非常普遍，一  
DATAMINE 軟件  
技術加入其系統中。  
樹 (DECISION TREE)。  
應用廣泛，所以請  
作、優點和缺點。  
適合用的軟件，而且在應用這  
候，能可選擇就  
AMING 的能力  
請決策樹，顧名思義就是將整  
的流程，把其中可  
件，有系統地以樹  
。決策樹是由樹根、樹幹和樹  
成 (見圖一)。樹  
樹葉是終點，樹  
解決策點)。

# Essence of DataMining

- Data and knowledge represented with simple mechanisms (typically, HTML) and without metadata (data about data).
- Consequently, relatively complex algorithms have to be used (complexity migrated into the retrieval request time).
- In return, low complexity at system design time!

# Essence of Semantic Web



- Data and knowledge represented with complex mechanisms (typically XML) and with plenty of metadata (a byte of data may be accompanied with a megabyte of metadata).
- Consequently, relatively simple algorithms can be used (low complexity at the retrieval request time).
- However, large metadata design and maintenance complexity at system design time.



# Major Knowledge Retrieval Algorithms (for DataMining)



- Neural Networks
- Decision Trees
- Rule Induction
- Memory Based Reasoning, etc...

Consequently, the stress is on algorithms!

# Major Metadata Handling Tools (for Semantic Web)



- XML
- RDF
- Ontology Languages
- Verification (Logic + Trust) Efforts in Progress

Consequently, the stress is on tools!



DATAMINING 軟件時，又需要輸入...  
能啓動工序，而所得到的結果又沒有甚麼...  
導致這個原因多數是用家不明白每個工序...  
和參數的用途。其實輸入參數有助引導工...  
有限的時間裏尋找有用的資料。自動化的...  
MINING 是在於你能否了解工件，配合參...  
彈 DATAMINING 的能力，從而得到真正

# Issues in Data Mining Infrastructure



適合用的軟件，而且在應用這...  
時候，能可避重就輕，大大發...  
AMINING 的功力。>  
請決策樹，顧名思義就是將整...  
的流程，把其中可能牽涉到的...  
件，有系統地以樹狀的結構呈...  
。決策樹是由樹根、樹節和樹...  
成（見圖一）。樹根是分析的...  
根，是終點、樹節是岔路...  
（決策點）。

事例設定決策樹  
[在用一...  
!，資料取材於香港賽馬紀錄，...  
用其中一場參賽馬（共十四...  
過往賽蹟（見表一）。現在來...  
!策樹的運作。首先決策樹演式...  
!放在起點的樹根，當用家確定...  
!析的特徵後，演式把其餘的特...  
!分析，對每一個互影響力的特



**Authors:**  
**Nemanja Jovanovic, [nemko@acm.org](mailto:nemko@acm.org)**  
**Valentina Milenkovic, [tina@eunet.yu](mailto:tina@eunet.yu)**  
**Veljko Milutinovic, [vm@etf.bg.ac.yu](mailto:vm@etf.bg.ac.yu)**

<http://galeb.etf.bg.ac.yu/vm>

DATAMINE 軟件時，又需要...  
能啓動工序，而所得到的結果又沒有...  
導致這個原因多數是用家不明白每個工序...  
和參數的用途。其實輸入參數有助引導工...  
有限的時間裏尋找有用的資料。自動化的...  
MINING 在於你能否了解工序，配合參...  
揮 DATAMINE 的能力，從而得到真正...  
的資料。



# Semantic Web

Ivana Vujovic (ile@eunet.yu)  
Erich Neuhold (neuhold@ipsi.fhg.de)  
Peter Fankhauser (fankhaus@ipsi.fhg.de)  
Claudia Niederée (niederee@ipsi.fhg.de)  
Veljko Milutinovic (vm@etf.bg.ac.yu)

<http://galeb.etf.bg.ac.yu/vm>



# Data Mining in the Nutshell



- Uncovering the hidden knowledge
- Huge n-p complete search space
- Multidimensional interface



# A Problem ...

## You are a marketing manager for a cellular phone company

- **Problem: Churn is too high**
- **Turnover (after contract expires) is 40%**
- **Customers receive free phone (cost 125\$) with contract**
- **You pay a sales commission of 250\$ per contract**
- **Giving a new telephone to everyone whose contract is expiring is very expensive (as well as wasteful)**
- **Bringing back a customer after quitting is both difficult and expensive**



# ... A Solution



- Three months before a contract expires, **predict** which customers will leave
- If you want to keep a customer that is predicted to churn, offer them a new phone
- The ones that are not predicted to churn need no attention
- If you don't want to keep the customer, do nothing
- How can you predict future behavior?



- Tarot Cards?
- Magic Ball?
- Data Mining?



# Still Skeptical?



如何提出分裂目標

空成選擇數據種類，可以開始選

來？方法是有很多，比較簡單的方法

是用分裂後的數據數量，當分出來的

廣泛，和5%的數據，

比這結果，再選出



# The Definition

## The automated extraction of predictive information from (large) databases

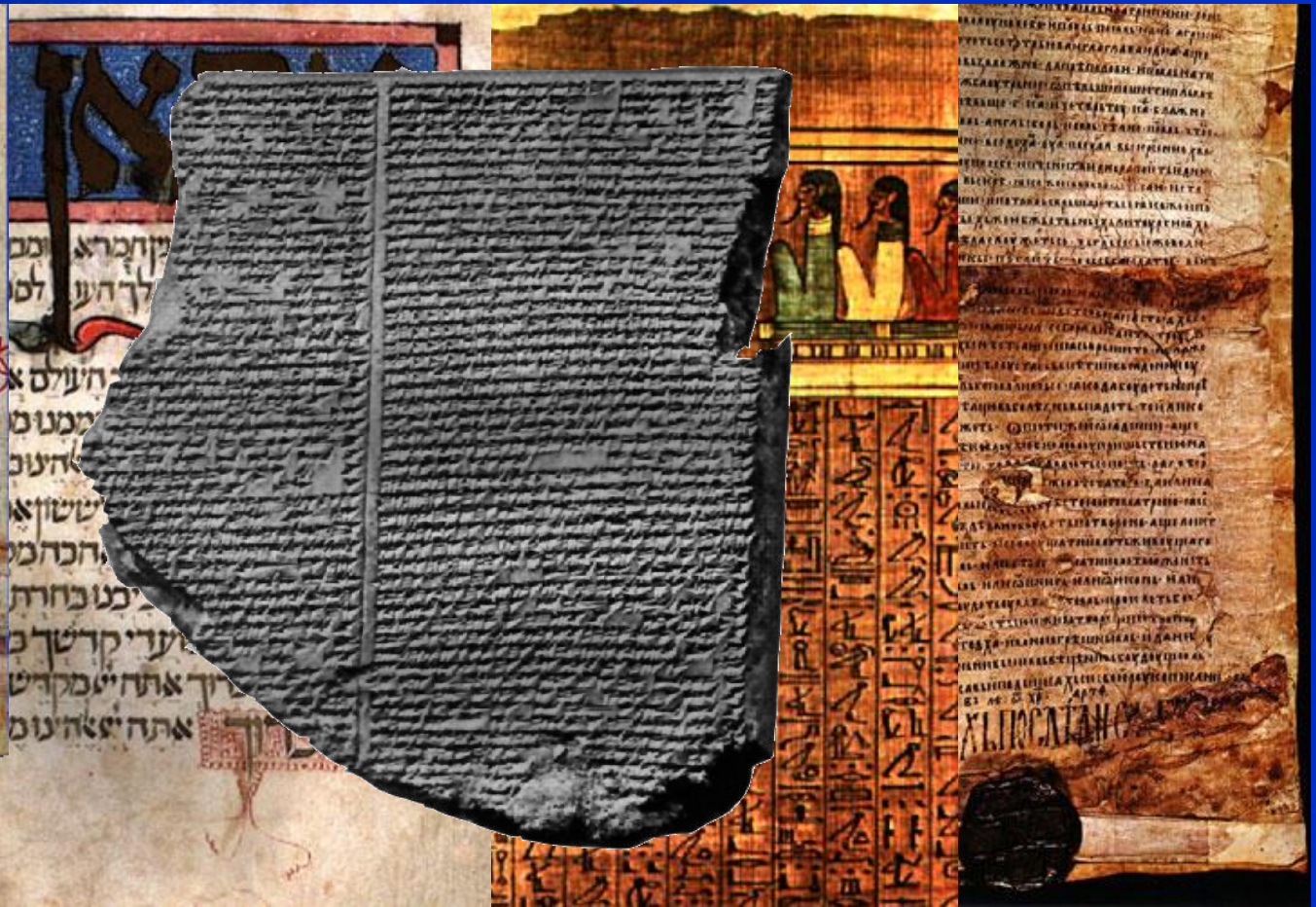
- Automated
- Extraction
- Predictive
- Databases

### 如何提出分裂目標

完成選擇數據種類，可以開始選



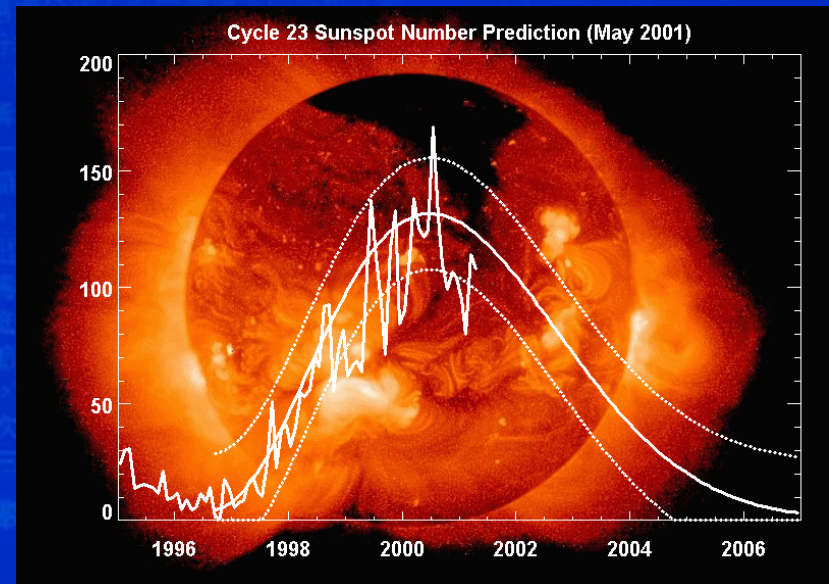
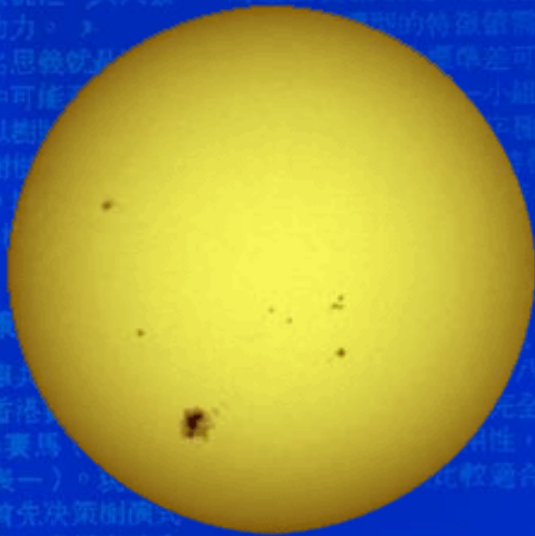
# History of Data Mining





# Repetition in Solar Activity

- 1613 – Galileo Galilei
- 1859 – Heinrich Schwabe



# The Return of the Halley Comet

## Edmund Halley (1656 - 1742)



1531

1607

1682

239 BC

1910

1986

2061 ???



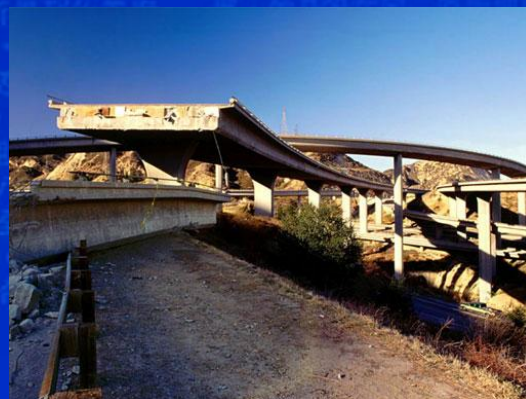
# Data Mining is Not

- Data warehousing
- Ad-hoc query/reporting
- Online Analytical Processing (OLAP)
- Data visualization



# Data Mining is

- Automated extraction of predictive information from various data sources
- Powerful technology with great potential to help users focus on the most important information stored in data warehouses or streamed through communication lines





# Focus of this Presentation



- **Data Mining problem types**

- **Data Mining models and algorithms**

- **Efficient Data Mining**

- **Available software**



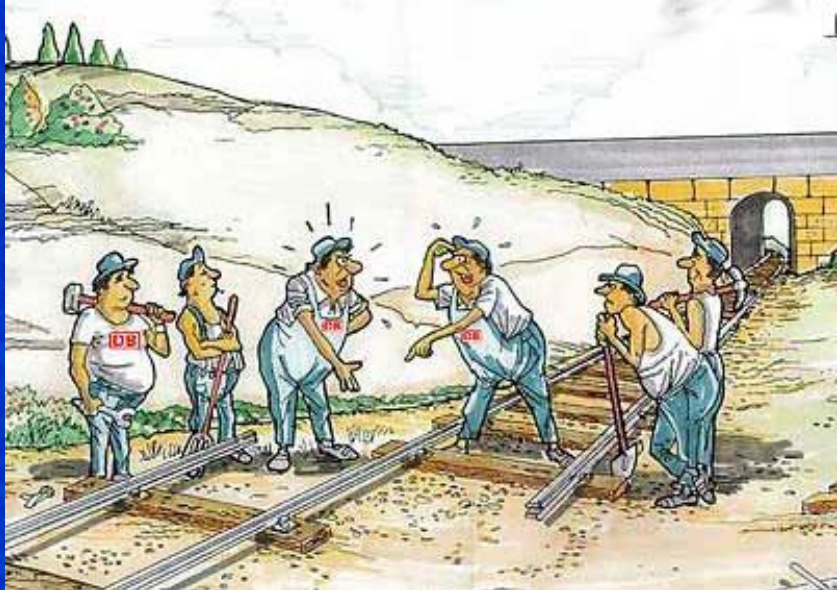


# Data Mining Problem Types



6 types

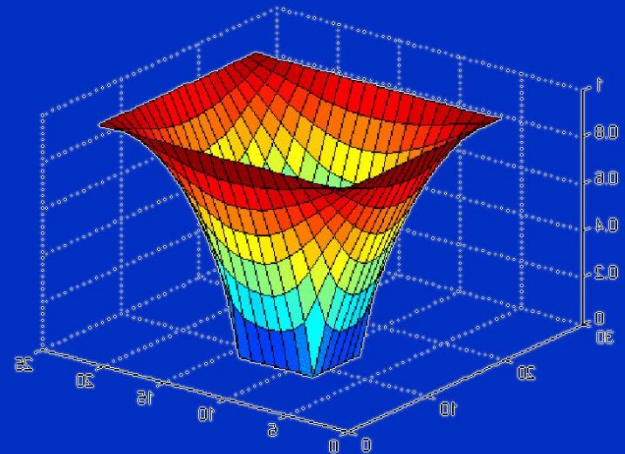
Often a combination solves the problem



# Data Description and Summarization



- Aims at concise description of data characteristics
- Lower end of scale of problem types
- Provides the user an overview of the data structure
- Typically a sub goal





# Segmentation

- Separates the data into interesting and meaningful subgroups or classes
- Manual or (semi)automatic
- A problem for itself or just a step in solving a problem



# Classification



- **Assumption: existence of objects with characteristics that belong to different classes**
- **Building classification models which assign correct labels in advance**
- **Exists in wide range of various application**
- **Segmentation can provide labels or restrict data sets**

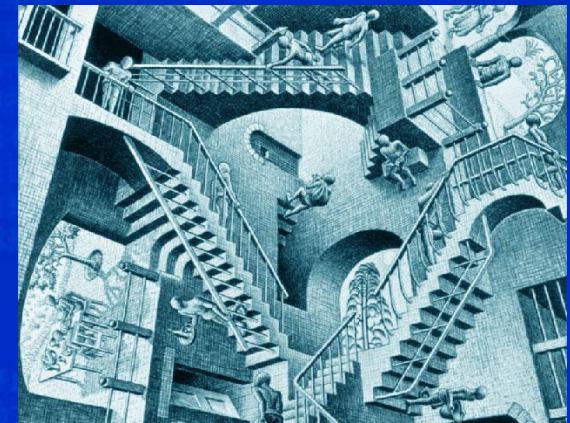




# Concept Description



- **Understandable description of concepts or classes**
- **Close connection to both segmentation and classification**
- **Similarity and differences to classification**



# Prediction (Regression)

- Finds the numerical value of the target attribute for unseen objects
- Similar to classification - difference: discrete becomes continuous





# Dependency Analysis



- Finding the model that describes significant dependences between data items or events
- Prediction of value of a data item
- Special case: associations



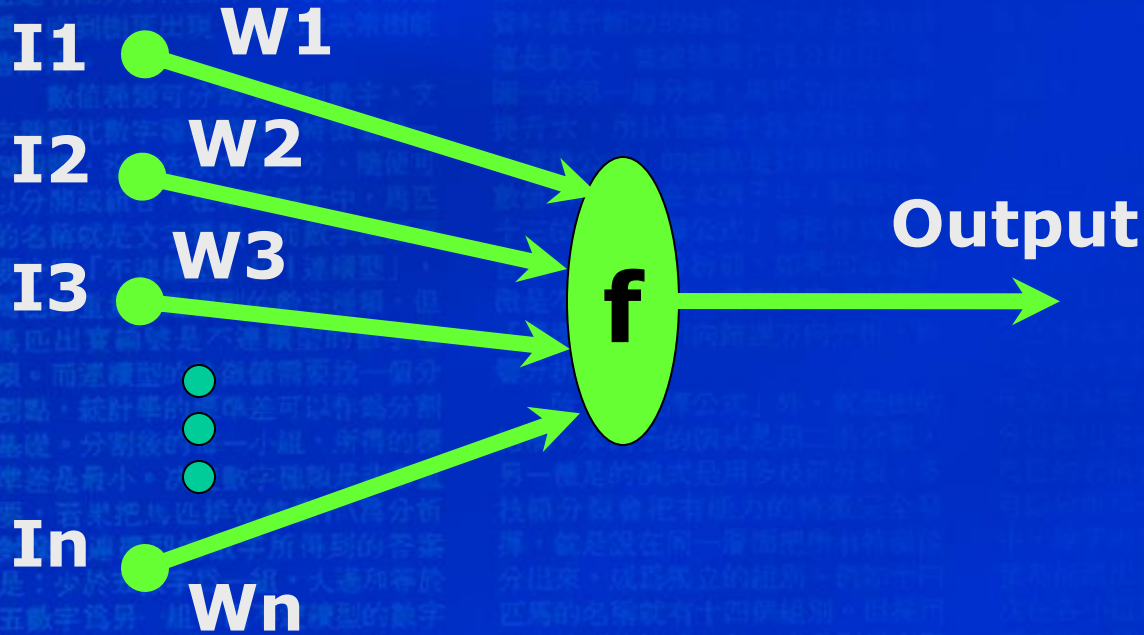


# Neural Networks

- Characterizes processed data with single numeric value
- Efficient modeling of large and complex problems
- Based on biological structures
- Network consists of neurons grouped into layers



# Neuron Functionality



$$\text{Output} = f(W_1 * I_1, W_2 * I_2, \dots, W_n * I_n)$$



# Training Neural Networks



# Decision Trees



- A way of representing a series of rules that lead to a class or value
- Iterative splitting of data into discrete groups maximizing distance between them at each split
- Classification trees and regression trees
- Univariate splits and multivariate splits
- Unlimited growth and stopping rules
- CHAID, CHART, Quest, C5.0





# Decision Trees

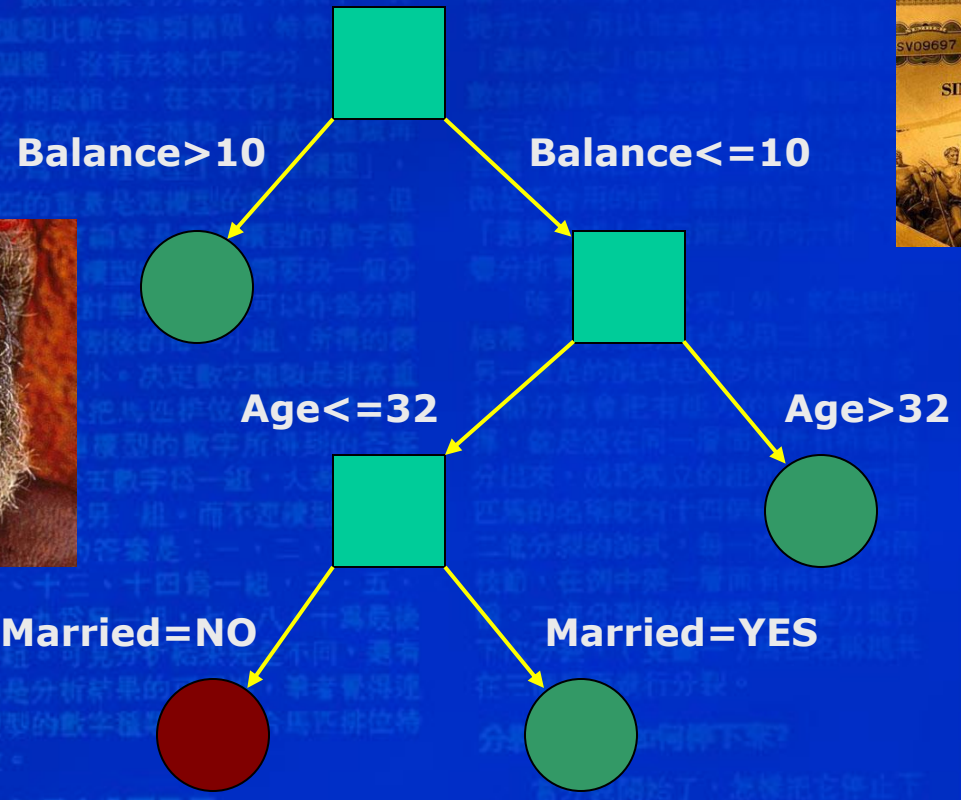


一輪的分裂中被選用。當分裂出來的子集過有能力的特選時，分裂會不斷循環，直到樹葉出現，樹狀的決策樹就會形成。

數值種類可分為文字和數字。文字種類比數字種類簡單，特別是像價錢，沒有先後次序之分，可以分開或組合。在本文例子中，要分析的是連環型的數字，但「連環型」的數字種類，在分類時，可以作為分割的標準。所以，所謂的「連環型」的數字，決定數字是取非零或零也拆成兩個子集。而「連環型」的數字所帶有的五數字為一組，入到另外一組，而不考慮到另一組。而「連環型」的數字所帶有的五數字為一組，入到另外一組，而不考慮到另一組。

如何提出分裂目標

完成選擇數據種類，可以開始過









# Rule Induction



- Method of deriving a set of rules to classify cases
- Creates independent rules that are unlikely to form a tree
- Rules may not cover all possible situations
- Rules may sometimes conflict in a prediction



# Rule Induction

**If balance > 100.000  
then confidence = HIGH & weight = 1.7**

**If balance > 25.000 and  
status = married  
then confidence = HIGH & weight = 2.3**

**If balance < 40.000  
then confidence = LOW & weight = 1.9**





# K-nearest Neighbor and Memory-Based Reasoning (MBR)



- **Usage of knowledge of previously solved similar problems in solving the new problem**

- **Assigning the class to the group where most of the k-“neighbors” belong**

- **First step – finding the suitable measure for distance between attributes in the data**

- **How far is black from green?**

- **+ Easy handling of non-standard data types**

- **- Huge models**

# K-nearest Neighbor and Memory-Based Reasoning (MBR)



如何提出分裂目標



# Data Mining Models and Algorithms



- Many other available models and algorithms
  - Logistic regression
  - Discriminant analysis
  - Generalized Adaptive Models (GAM)
  - Genetic algorithms
  - Etc...
- Many application specific variations of known models
- Final implementation usually involves several techniques
- Selection of solution that match best results

DATAMINING 軟件時，又需要輸入一些數據，才能啓動工序，而所得到的結果又沒有甚麼導致這個原因多數是用家不明白每個工序和參數的用途。其實輸入參數有助引導工友有限的時間裏尋找有用的資料。自動化的 DATAMINING 是在於你能否了解工序，配合發揮 DATAMINING 的能力，從而得到真正

馬名	馬主	訓練師	出賽日期	出賽地點	出賽場次	出賽成績
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王
馬中王	馬中王	馬中王	馬中王	馬中王	馬中王	馬中王

# Efficient Data Mining



適合用的軟件，而且在應用這時候，能可避重就輕，大大發揮 DATAMINING 的功力。

讀決策樹，顧名思義就是將整的流程，把其中可能牽涉到的條件，有系統地以樹狀的結構呈現。決策樹是由樹根、樹節和樹枝成（見圖一）。樹根是分析的起點，樹節是分割點，樹枝是分割路（決策點）。

## 操作例設定決策樹

「在用一個數據庫共三百三十二宗，資料取材於香港賽馬紀錄，我用其中一場賽馬（共十四場）過往賽蹟（見表一）。現在來談決策樹的運作。首先決策樹演式放在起點的樹根，當用家確定分析的特徵後，演式把其餘的特徵分析，產生一個互影響力的特

馬匹的數量是連續型的數字種類，但馬匹出賽編號是不連續型的數字種類。而連續型的特徵值需要找一個分割點，統計學的標準差可以作為分割基礎。分割後的每一小組，所得的標準差是最小。決定數字種類是非常重要的，若果把馬匹排位特徵作為分析對象，連續型的數字所得到的答案是：少於五數字為一組，大過和等於五數字為另一組。而不連續型的數字所得到的答案是：一、二、四、十一、十二、十四為一組，三、五、六、九為另一組，七、八、十為最後一組。可見分析結果完全不同，還有的是分析結果的可用性，筆者覺得連續型的數字種類比較適合馬匹排位特徵。

## 如何選出分裂目標

完成選擇數據種類，可以開始選擇分裂目標。演式把演式公式套用在每個特徵，計算每個特徵的公

而是不合用的話，請到「選擇公式」引向更深層次的內容。本層分析實例！

除了「選擇公式」外，較常用的結構。本例子的演式是用二邊分裂，另一種是節的演式是用多枝和分裂。多枝和分裂會把有能力的特徵完全捨棄，就是說在同一層面把所有特徵捨分出來，成為那邊的組別，而如馬匹馬的名稱就有十四個組別。但採用二邊分裂的演式，每一次分裂只有兩枝節，在例中第一層面有兩枝節是名稱。二邊分裂後的特徵還有能力進行下次分裂。（見圖一）馬匹名稱總共在三個層面進行分裂。

## 分裂開始後如何停下來？

當分裂開始了，怎樣把它停止下來？方法是有很多，比較簡單的方法是用分裂後的數據數量。當分出來的組別越多，每一個組別的數據數量越少，則果在分析前可訂下一層面，

而「選擇公式」引向更深層次的內容。本層分析實例！

除了「選擇公式」外，較常用的結構。本例子的演式是用二邊分裂，另一種是節的演式是用多枝和分裂。多枝和分裂會把有能力的特徵完全捨棄，就是說在同一層面把所有特徵捨分出來，成為那邊的組別，而如馬匹馬的名稱就有十四個組別。但採用二邊分裂的演式，每一次分裂只有兩枝節，在例中第一層面有兩枝節是名稱。二邊分裂後的特徵還有能力進行下次分裂。（見圖一）馬匹名稱總共在三個層面進行分裂。

## 這種分析會不會「死機」？

「馬匹出賽編號」是「馬匹名稱」的互影響力是低的。因為馬匹在分析前訂下「馬匹名稱」的條件，即分析馬匹出賽編號4、11、12的數目，這時就只剩下十個檔案，再加上有資料沒有馬匹名稱的檔案，所以分析到最後，





# DM Process Model



- **5A – used by SPSS Clementine**  
(Assess, Access, Analyze, Act and Automate)
- **SEMMA – used by SAS Enterprise Miner**  
(Sample, Explore, Modify, Model and Assess)
- **CRISP-DM – tends to become a standard**



DATAMINE 軟件時，又需要輸入一些數據，才能啟動工序，而所得到的結果又沒有直接導致這個原因多數是用家不明白每個工序和參數的用途。其實輸入參數有助引導工



# CRISP - DM

- Cross-Industry Standard for DM
- Conceived in 1996 by three companies:

DAIMLERCHRYSLER



### 如何提出分裂目標

完成選擇數據種類，可以開始選

# CRISP – DM methodology



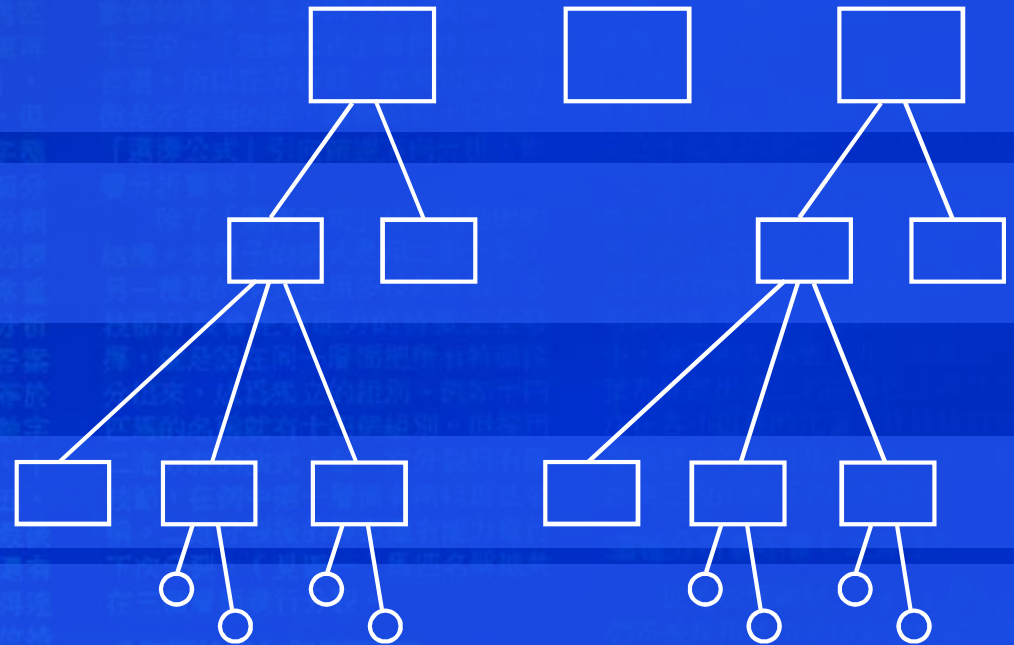
## Four level breakdown of the CRISP-DM methodology:

Phases

Generic Tasks

Specialized Tasks

Process Instances





# Mapping generic models to specialized models



- **Analyze the specific context**
- **Remove any details not applicable to the context**
- **Add any details specific to the context**
- **Specialize generic context according to concrete characteristic of the context**
- **Possibly rename generic contents to provide more explicit meanings**

# Generalized and Specialized Cooking



- Preparing food on your own
- **Raw steak with vegetables?**

- Find the recipe for that meal
- Check the Cookbook or call mom
- Gather the ingredients
- Defrost the meat (if you had it in the fridge)
- Prepare the meal
- Buy missing ingredients
- Enjoy your food

- Clean up (later)
- Cook (later)



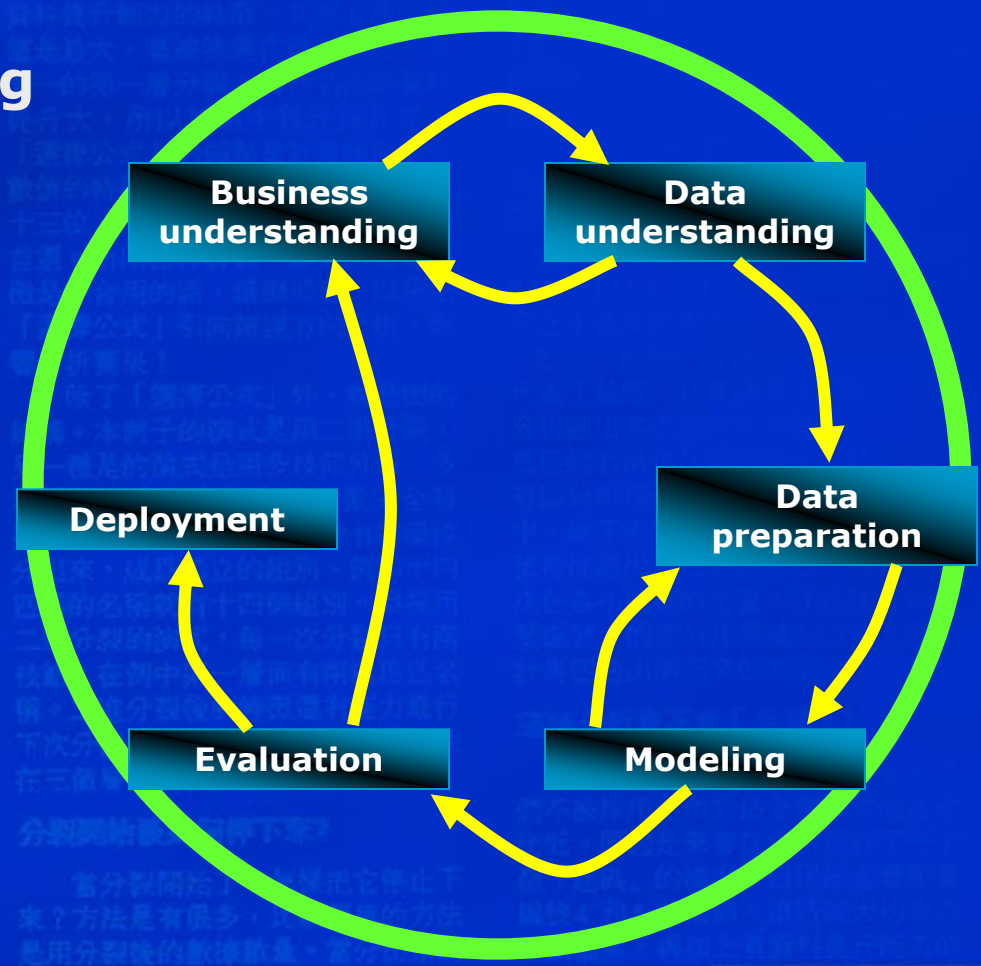
the dishes



# CRISP - DM model



- Business understanding
- Data understanding
- Data preparation
- Modeling
- Evaluation
- Deployment



# Business Understanding



- Determine business objectives
- Assess situation
- Determine data mining goals
- Produce project plan





# Data Understanding

- Collect initial data
- Describe data
- Explore data
- Verify data quality



# Data Preparation

- **Select data**
- **Clean data**
- **Construct data**
- **Integrate data**
- **Format data**





# Modeling

- Select modeling technique
- Generate test design
- Build model
- Assess model

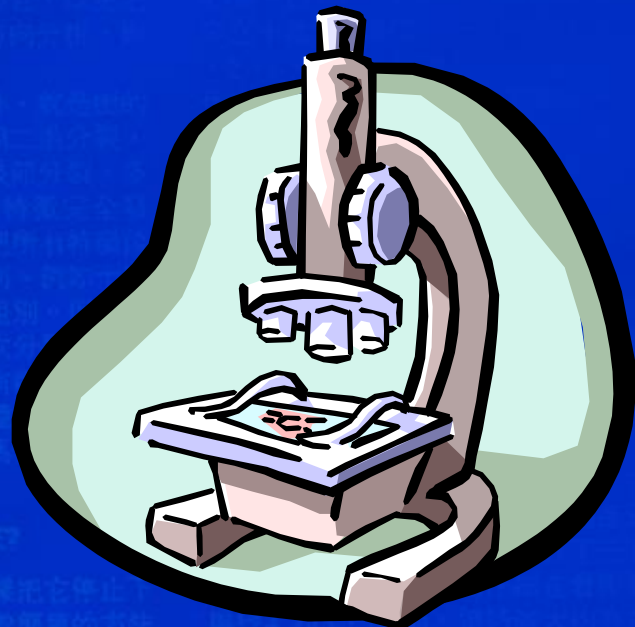


# Evaluation



results = models + findings

- Evaluate results
- Review process
- Determine next steps





# Deployment

- Plan deployment
- Plan monitoring and maintenance
- Produce final report
- Review project



DATA MINING 軟件時，又需要輸入  
能啓動工序，而所得到的結果又沒有直接  
導致這個原因多數是用家不明白每個工序  
和參數的用途。其實輸入參數有助引導工  
在有限的時間裏尋找有用的資料。自動化的  
MINING 在於你能否了解工序，配合參  
DATA MINING 的能力，從而得到真正  
的資料。



# At Last...

DATA MINING 至今讀者  
看到一些字眼，例如分類、拼  
法條等。它們都是在  
MINING 技術中出來的，可以  
技術的應用是非常普遍，一半  
DATAMINE 軟件開發公司，  
技術加入其系統中，這技術就  
樹 (DECISION TREE)。由  
應用廣泛，所以清楚掌握決策  
作、優點和缺點，不但可以為  
擇合用的軟件，而且在應用這  
時候，能可避重就輕，大大發  
DATA MINING 的功力。

請決策樹，顧名思義就是將整  
的流程，把其中可能牽涉到的  
件，有系統地以樹狀的結構呈  
。決策樹是由樹根、樹節和樹  
成 (見圖一)。樹根是分析的  
樹葉是終點，樹節是分支路  
 (決策點)。

## 為作例設定決策樹

現在用一個數據庫共三百三十二  
，資料取材於香港賽馬紀錄，  
用其中一場賽馬 (共十四  
過往賽蹟 (見表一)。現在來  
決策樹的運作。首先決策樹演式  
放在起點的樹根，當用家確定  
分析的特徵後，演式把其餘的特  
分析，變成一個具影響力的特



特征是大小、穩定數、穩定  
要，若果把其  
對像，連續型  
是：少於五數  
五數字為另  
所得到的答  
一、十二、十  
六、九為另  
一組。可見分  
的是分析結果  
類型的數字種  
數。

## 如何提出分

變成選擇數據庫，可以用指定  
用分析特徵的數據點，為分析  
來的







# Comparison of fourteen DM tools



- The Decision Tree products were
  - *CART*
  - *Scenario*
  - *See5*
  - *S-Plus*
- The Rule Induction tools were
  - *WizWhy*
  - *DataMind*
  - *DMSK*
- Neural Networks were built from three programs
  - *NeuroShell2*
  - *PcOLPARS*
  - *PRW*
- The Polynomial Network tools were
  - *ModelQuest Expert*
  - *Gnosis*
  - a module of *NeuroShell2*
  - *KnowledgeMiner*



# Criteria for evaluating DM tools



A list of 20 criteria for evaluating DM tools, put into 4 categories:

- *Capability* measures what a desktop tool can do, and how well it does it
  - Handles missing data
  - Considers misclassification costs
  - Allows data transformations
  - Quality of testing options
  - Has programming language
  - Provides useful output reports
  - Visualisation



# Criteria for evaluating DM tools



- *Learnability/Usability* shows how easy a tool is to learn and use:

- Tutorials
- Wizards
- Easy to learn
- User's manual
- Online help
- Interface





# Criteria for evaluating DM tools

- *Interoperability* shows a tool's ability to interface with other computer applications

- Importing data
- Exporting data
- Links to other applications

- *Flexibility*

- Model adjustment flexibility
- Customizable work environment
- Ability to write or change code



# A classification of data sets

- Pima Indians Diabetes data set

- 768 cases of Native American women from the Pima tribe some of whom are diabetic, most of whom are not
- 8 attributes plus the binary class variable for diabetes per instance

- Wisconsin Breast Cancer data set

- 699 instances of breast tumors some of which are malignant, most of which are benign
- 10 attributes plus the binary malignancy variable per case

- The Forensic Glass Identification data set

- 214 instances of glass collected during crime investigations
- 10 attributes plus the multi-class output variable per instance

- Moon Cannon data set

- 300 solutions to the equation:  
$$x = 2v^2 \sin(g)\cos(g)/g$$
- the data were generated without adding noise





# Evaluation of fourteen DM tools

Technology	Tool	Capability	Learnability/ Usability	Interoperability	Flexibility	Accuracy	Overall (equal weights)	Price (US\$)
Tree	<b>CART</b>	+	✓	-	✓+	+	✓+	995
	<b>Scenario</b>	✓-	+	+	-	--	✓	695
	<b>See5</b>	✓	✓-	✓	✓-	+	✓	440
	<b>S-Plus</b>	+	✓-	++	+	+	+	1,795
	Tree Average*	✓	✓	✓+	✓	✓+	✓+	Median = 845
Rule	<b>WizWhy</b>	✓	✓+	✓	✓-	-	✓	4,000
	<b>DataMind</b>	✓+	++	+	✓-	✓	✓+	25,000
	<b>DMSK</b>	-	--	✓-	-	+	-	75
	Rule Average*	✓	✓	✓+	-	✓	✓	Median = 4,000
Neural	<b>NeuroShell 2</b>	-	✓	-	-	++	✓-	395
	<b>PcOLPARS</b>	✓-	-	-	✓-	✓	✓-	495
	<b>PRW</b>	✓+	+	++	✓	++	+	10,000
	Neural Average*	✓-	✓	✓	✓-	+	✓-	Median = 495
Poly Net	<b>MQ Expert</b>	+	✓	✓	✓+	+	✓+	5,950
	<b>NeuroShell 2</b>	✓-	✓	✓	✓	+	✓	495
	<b>Gnosis</b>	✓-	✓	--	✓-	++	✓-	4,900
	<b>KnowledgeMiner</b>	-	-	-	✓-	+	-	100
	Poly Net Average*	✓-	✓-	-	✓	+	✓-	Median = 2,698
Overall Average*	✓	✓	✓	✓	✓-	✓+	✓	Median = 845





# WWW.NBA.COM

NBA.com - Microsoft Internet Explorer

File Edit View Favorites Tools Help

← Back → Forward Stop Home Search Favorites History

Address [http://www.nba.com/statistics/default\\_finals\\_leaders/index.html?query=LeagueLeadersPPGQu](http://www.nba.com/statistics/default_finals_leaders/index.html?query=LeagueLeadersPPGQu)

Click here for a free download!

Global FANTA

## STATISTICS

STAT SEARCH

Select a Topic: **Finals League Leaders** Select a Category: **Points Per Game** GO

### 2000-01 FINALS POINTS PER GAME

PLAYER	G	FG	FT	P	PG
1. Allen Iverson ( Philadelphia 76ers)	5	66	35	178	35.6
2. Shaquille O'Neal ( Los Angeles Lakers)	5	63	39	165	33.0
3. Kobe Bryant ( Los Angeles Lakers)	5	44	32	123	24.6
4. Dikembe Mutombo ( Philadelphia 76ers)	5	33	18	84	16.8
5. Eric Snow ( Philadelphia 76ers)	5	22	19	63	12.6
6. Derek Fisher ( Los Angeles Lakers)	5	17	5	49	9.8
6. Rick Fox ( Los Angeles Lakers)	5	15	12	49	9.8
8. Robert Horry ( Los Angeles Lakers)	5	14	6	42	8.4
9. Aaron McKie ( Philadelphia 76ers)	5	15	6	40	8.0
10. Tyrone Hill ( Philadelphia 76ers)	5	13	7	33	6.6

Done Internet

DATAMINE 軟件時，又需要輸入一些參數才能啓動工序，而所得到的結果又沒有意義導致這個原因多數是用家不明白每個工序和參數的用途。其實輸入參數有助引導工人在有限的時間裏尋找有用的資料。自動化的 DATAMINE 在於你能否了解工序，配合參數發揮 DATAMINE 的能力，從而得到真正



# Se7en



~~GLUTTONY~~

~~GREED~~

~~SLOTH~~

~~ENVY~~

~~WRATH~~

~~PRIDE~~

~~LUST~~

Seven deadly sins. Seven ways to die.

# SEVEN





