

Introduction to Machine Learning

Instructor: Kaiqi Zhao

Overview

- What is machine learning?
- Data mining, AI, Deep learning
- Applications

Age of Big Data

- We are collecting and storing data at unprecedented rate
 - Youtube, Facebook, news
 - Credit card transactions and Amazon purchases
 - Transportation data (Google Maps, Uber, Auckland Transport)
 - Gene expression data, protein interaction assays, chemistry pathways

Age of Big Data

- What can you do with this data?
 - Too much data to search through it manually?
- But there is valuable information in the data
 - How can we use it for fun, profit, and/or the greater good?
- **Machine Learning** and **Data Mining** are key tools we use to make sense of large data sets!

What is Machine Learning?

- “*Machine Learning is the field of study that gives the computer the ability to learn without being explicitly programmed.*” – Arthur Samuel (1959)
- “*A computer is said to learn from experience E with respect to some class of task T and performance measure P, if its performance at tasks in T, as measured by P, improved with experience E*” – Tom Mitchell



Arthur Samuel



Example Problem – Self-Driving Cars

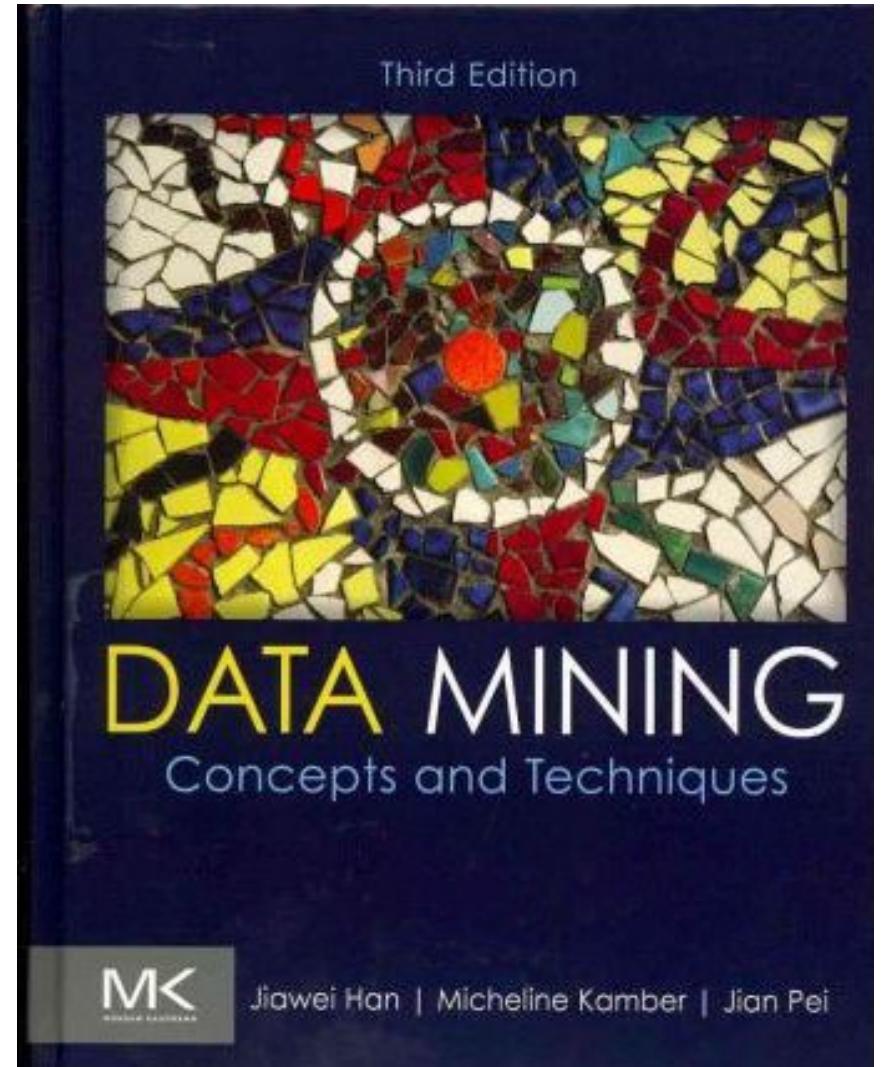
- T : Driving on public highway using vision sensors
- P : Average distance traveled before an error / crash
- E : A sequence of images and steering commands recorded while observing a human driver



Image source: <https://www.forbes.com/>

Data Mining

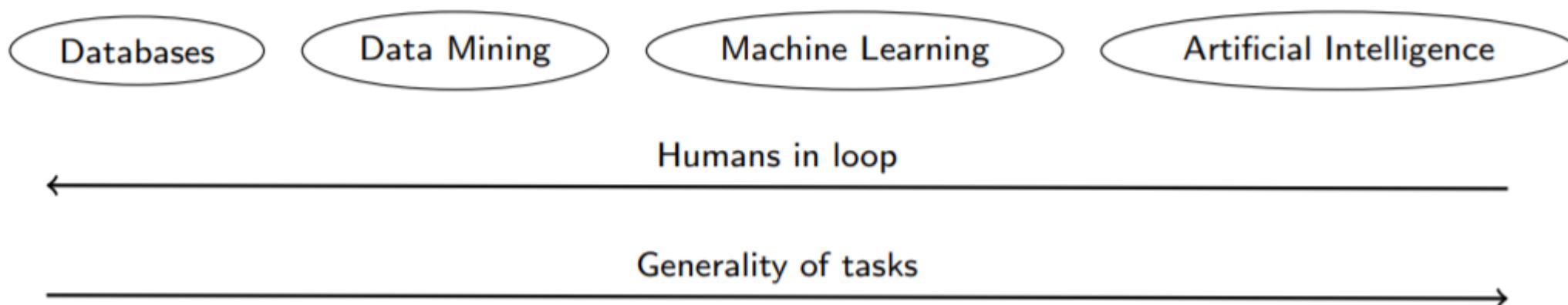
- “*Data mining refers to extracting or mining knowledge from large amounts of data*” – Jiawei Han





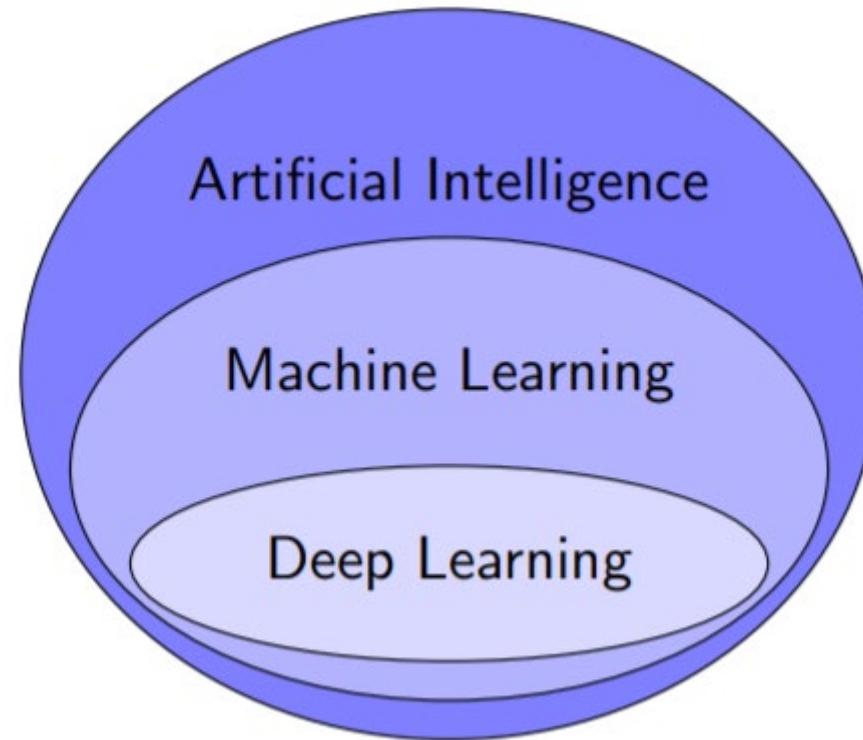
Machine Learning vs. Data Mining

- Data Mining and Machine Learning are very similar:
 - Data Mining often viewed as closer to databases
 - Machine Learning often viewed as closer to AI
- Both are similar to statistics, but more emphasis on
 - Large data sets and computation
 - Predictions (instead of descriptions)
 - Flexible models (that work on many problems)



AI and Deep Learning

- Traditionally, we've viewed ML as a subset of AI
- Deep learning is a subclass of machine learning



Applications

- Machine translation
- Speech recognition
- Face recognition
- Image generation
- Artistic style transfer
- Recommendations
- Object detection
- Optical character recognition
- Sports analytics
- Medical imaging
- Self-driving cars
- Image annotation
- Discovering new cancer subtypes
- Protein folding prediction
- Drug discovery
- Chess / Go / Starcraft / ...

Machine translation

Google Translate Turn off instant translation

English Spanish Chinese Detect language ▾ English Spanish Chinese (Simplified) ▾ Translate

人工智能里程碑：新AlphaGo Zero横空出世，彻底摆脱人类
 打败目前人类围棋最高水平选手才过去不久，DeepMind团队又宣布了一个在AI历史上具有里程碑式意义的重磅消息：新版AlphaGo可以彻底摆脱人类的知识。

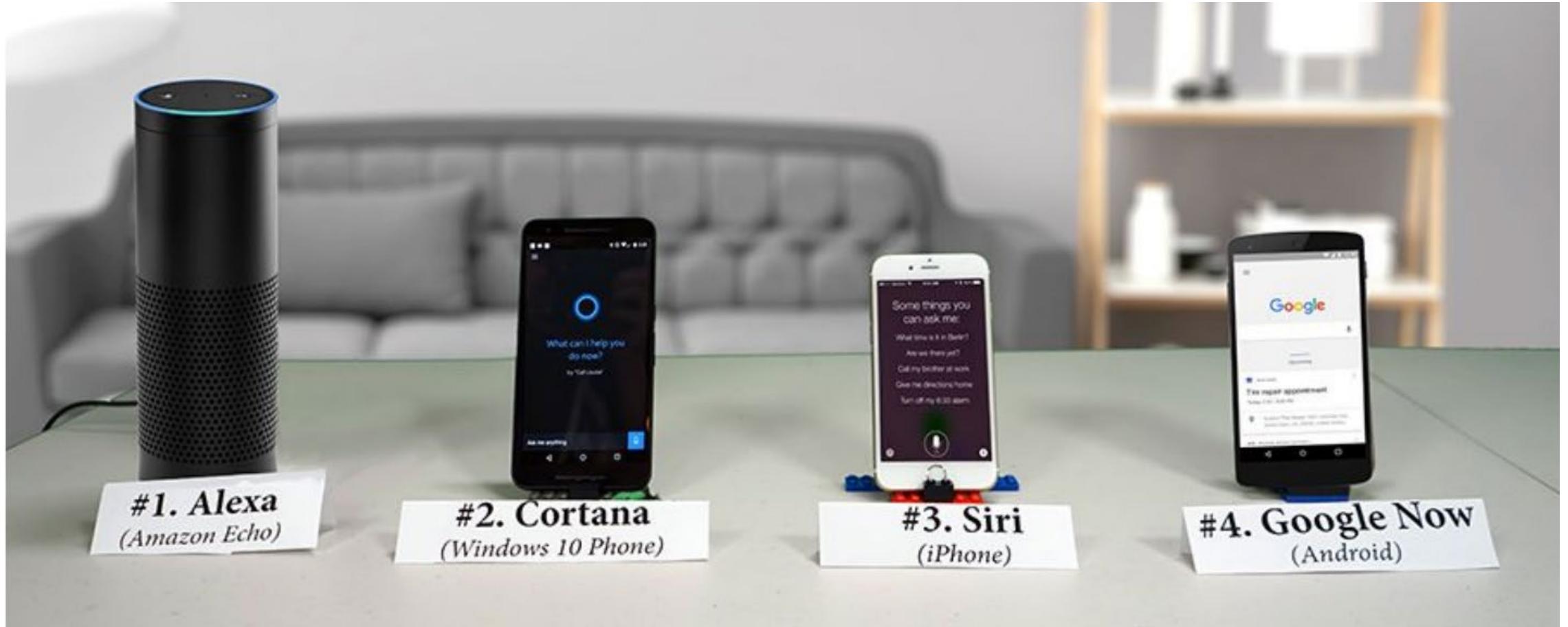
10月18日，谷歌人工智能团队DeepMind团队在期刊《Nature》上发表论文，宣布新版AlphaGo——AlphaGo Zero可以在没有人类指导的情况下学习。

也就是说，AlphaGo Zero就像人类初学者一样，只是了解比赛的规则和比赛的最终目标来自我学习。而作为首个击败人类围棋冠军的程序，之前版本的AlphaGo是通过分析人类围棋高手的数千棋局来学习，并在自我对弈中不断强化。

309/5000

Artificial intelligence milestone: the new AlphaGo Zero turned out, completely out of mankind
 The DeepMind team also announced a milestone message in the history of AI: the new AlphaGo can completely get rid of human knowledge.
 On October 18th, the Google artificial intelligence team DeepMind team published a paper in the journal Nature, announcing that the new AlphaGo - AlphaGo Zero could learn without human guidance.
 In other words, AlphaGo Zero is like a human beginner, just to understand the rules of the game and the ultimate goal of the game comes from my study. And as the first to defeat the human go championship program, the previous version of AlphaGo is through the analysis of human chess master thousands of chess to learn, and in self-chess constantly strengthen.

Speech recognition



Autonomous driving

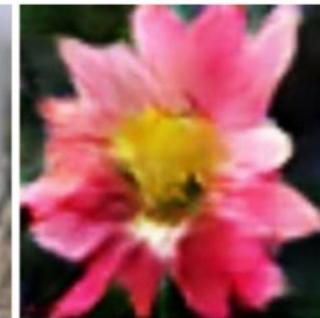
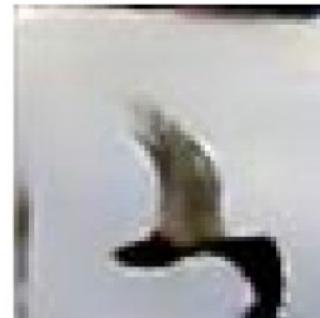


[Examples from Stanford CS221] 12

Image Generation

■ Text → Image

Generated
Image



Ground truth

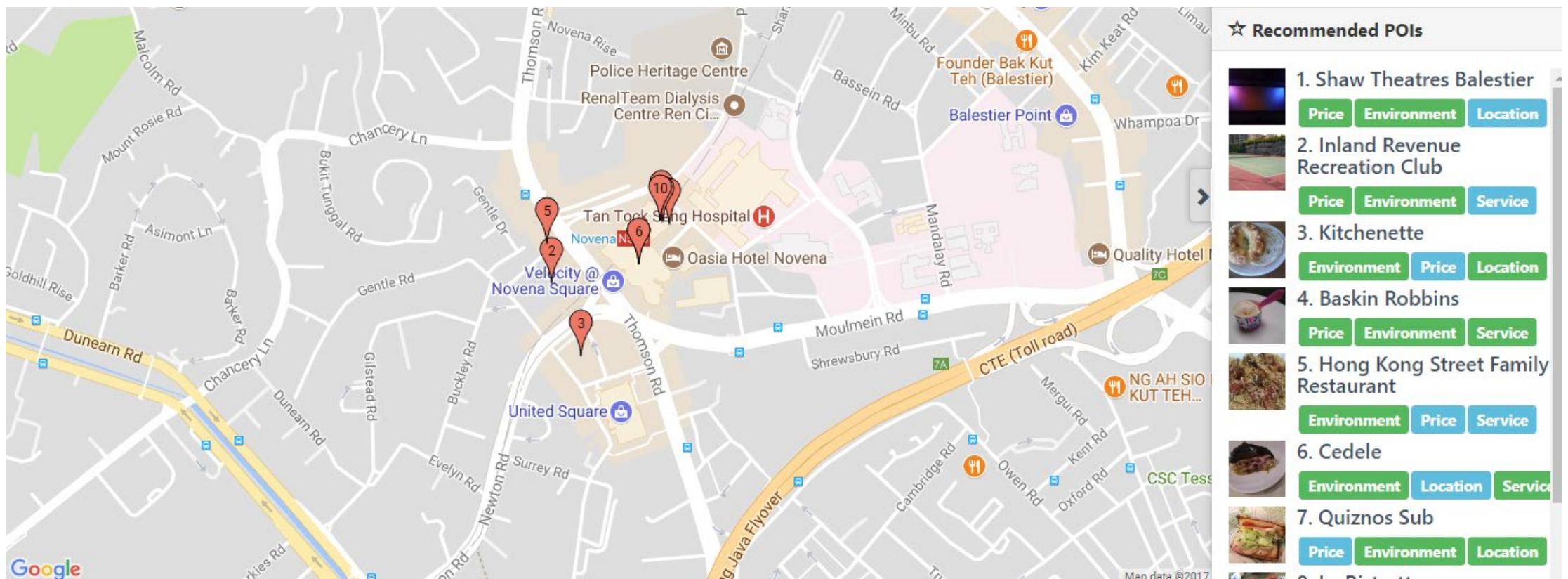


Artistic style transfer



Recommendation

- Recommend venues based on user's interests



Resources

- “Machine Learning” – Tom Mitchell
- “Data Mining: Concepts and Techniques” – Jiawei Han, Micheline Kamber, Jian Pei
- “Pattern Recognition and Machine Learning” – Christopher Bishop
- “Data Mining: Practical Machine Learning Tools and Techniques” – Ian Witten, Eibe Frank, Mark Hall, Christopher Pal