BITS Pilani

Pilani | Dubai | Goa | Hyderabad

S2-20_DSECLZC415 : Data Mining

(Lecture #17 - Data Mining Applications)

- *The slides presented here are obtained from the authors of the books and from various other contributors. I hereby acknowledge all the contributors for their material and inputs.*
- *I have added and modified a few slides to suit the requirements of the course.*

BITS Pilani

Pilani | Dubai | Goa | Hyderabad

Data Mining

Data Mining Applications

innovate achieve lead

BITS Pilani

Pilani | Dubai | Goa | Hyderabad

9.1 Recommendation Systems

Source Courtesy: Some of the contents of this PPT are sourced from materials provided by publishers of prescribed books

Recommendation System

- System that produces individualized recommendations as output or has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options

Burke (2002)

- A personalized information filtering technology used to either predict whether a particular user will like a particular item (prediction problem) or to identify a set of *N* items that will be of interest to a certain user (top-*N* recommendation problem).

Deshpande and Karypis (2004)

- Recommender systems suggest items of interest to users based on their explicit and implicit preferences, the preferences of other users, and user and item attributes

Schein et al. (2005)

Why Recommender Systems?

- Physical delivery systems are characterized by a scarcity of resources. Brick-and-mortar stores have limited shelf space, and can show the customer only a small fraction of all the choices that exist.
  - A physical bookstore may have several thousand books on its shelves
  - It is not possible to tailor the store to each individual customer
  - The choice of what is made available is governed only by the aggregate numbers
  - Typically, a bookstore will display only the books that are most popular, and a newspaper will print only the articles it believes the most people will be interested in.

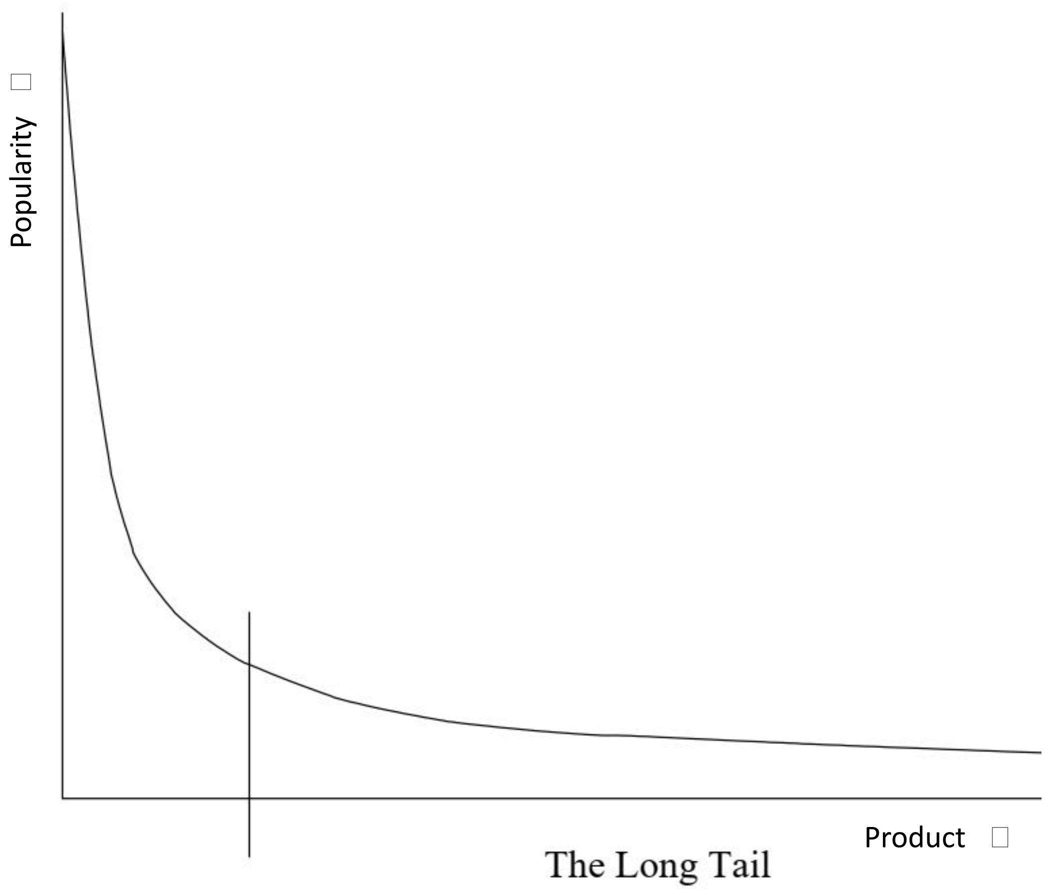
- On-line stores can make anything that exists available to the customer.
  - Largest ecommerce stores offer millions of books.
  - It is not possible to present all available items to the user, the way physical institution can. Neither can we expect users to have heard of each of the items they might like.

Why Recommender Systems?

The long tail: physical institutions can only provide what is popular,

while on-line institutions can make everything available

The long-tail phenomenon forces on-line institutions to recommend items to individual users.



Popular Recommendation Engines

Some popular recommendation engines which have been proved as highly profitable:

the Amazon.com,

the Netflix.com and

the Google.com recommendation engines

Amazon.com claims that 35 % of products sales result from recommendations.

About 66 % of movies rented in Netflix.com are recommended.

Google News Recommendations generate 38 % more click-throughs .

Recommender Systems for Location-based Social Networks by Panagiotis Symeonidis, Dimitrios Ntempos and Yannis Manolopoulos Springer © 2014

Into Thin Air and Touching the Void

An extreme example of how the long tail, together with a well designed recommendation system can influence events is the story told by Chris Anderson about a book called *Touching the Void*. This mountain-climbing book was not a big seller in its day, but many years after it was published, another book on the same topic, called *Into Thin Air* was published. Amazon's recommendation system noticed a few people who bought both books, and started recommending *Touching the Void* to people who bought, or were considering, *Into Thin Air*. Had there been no on-line bookseller, *Touching the Void* might never have been seen by potential buyers, but in the on-line world, *Touching the Void* eventually became very popular in its own right, in fact, more so than *Into Thin Air*.

Mining of Massive Datasets
By Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman

Recommendation Approaches

The approaches of recommender systems: collaborative filtering (CF), content-based Filtering (CB) and hybrid methods:

- ***Collaborative filtering*** algorithms recommend those items to the target user, that have been rated highly by other users with similar preferences and tastes. Websites that provide recommendations in the form, "Customers who bought item *i* also bought item *y*", typically fall under collaborative filtering approaches
- ***Content-based filtering*** uses the information derived from documents or item features (eg. terms or attributes). It uses a set of attributes, which describes the items and recommends other items similar to those that exist in the user's profile. This way, the cold start problem for new items and new users are alleviated, *provided* that users prefer items that are similar in content to those they have already chosen. However, the pitfall is that there is no diversity in the recommendations. That is, the user gets recommendations that are very familiar to her, since the recommended items are similar to those in her item profile
- ***Hybrid algorithms*** attempt to combine Collaborative filtering with Content-based filtering. The combination of content with rating data helps capture more effective correlations between users or items, which yields more accurate recommendations.

Recommendation System Types

Other types of recommender systems proposed (by Burke) in the literature :

- ***Demographic recommendation***, which classifies the users according to the attributes of their personal profile, and makes recommendations based on demographic classes
- ***Utility-based recommendation***, which makes suggestions based on a computation of the utility of each item for a user, for whom a utility function has to be stored
- ***Knowledge-based recommendation***, which suggests items based on logical inferences about user preferences. A knowledge representation (e.g. rules) about how an item meets a particular user need is necessary.

Utility matrix sample

	HP1	HP2	HP3	TW	SW1	SW2	SW3
A	4			5	1		
B	5	5	4				
C				2	4	5	
D		3					3

The movie names are HP1, HP2, and HP3 for Harry Potter I, II, and III, TW for Twilight, and SW1,
SW2, and SW3 for Star Wars episodes 1, 2, and 3. The users are represented
by capital letters A through D

A utility matrix representing ratings of movies on a 1–5 scale

Blanks represent the situation where the user has not rated the movie. In practice, the matrix would be even sparser

Recommendation Problem

- *Object of the decision*. That is, defining the object upon which the decision has to be made and the rationale of the recommendation decision
- *Family of criteria*. That is, the identification and modelling of a set of criteria that affect the recommendation decision, and which are exhaustive and non-redundant
- *Global preference model*. That is, the definition of the function that aggregates the marginal preferences upon each criterion into the global preference of the decision maker about each item
- *Decision support process*. That is, the study of the various categories and types of recommender systems that may be used to support the recommendation decision maker, in accordance to the results of the previous steps

Recommendation Capabilities

- *Choice*, which involves choosing one item from a set of candidates
- *Sorting*, which involves classifying items into pre-defined categories
- *Ranking*, which involves ranking items from the best one to the worst one
- *Description*, which involves describing all the items in terms of performance upon each criterion

Recommendation technique used in TEL

Collaborative filtering (CF) techniques				
Name	**Short description**	**Advantages**	**Disadvantages**	**Usefulness for TEL**
1. User-based CF	Users that rated the same item similarly probably have the same taste. Based on this assumption, this technique recommends unseen items already rated by similar users.	– No content analysis – Domain-independent – Quality improves over time – Bottom-up approach	– New user problem – New item problem – Popular taste – Scalability –Sparsity – Cold-start problem	– Benefits from experience –Allocates learners to groups (based on similar ratings)
2. Item-based CF	Focus on items, assuming that items rated similarly are probably similar. It recommends items with highest correlation (based on ratings to the items).	– No content analysis – Domain-independent – Quality improves over time – Bottom-up approach	– New item problem – Popular taste – Sparsity – Cold-start problem	– Benefits from experience
3. Stereotypes or demographics CF	Users with similar attributes are matched, then recommends items that are preferred by similar users (based on user data instead of ratings).	– No cold-start problem – Domain-independent	– Obtaining information –Insufficient information –Only popular taste –Obtaining metadata information –Maintenance ontology	–Allocates learners to groups –Benefits from experience –Recommendation from the beginning of the RS

TEL : Technology Enhanced Learning

Recommendation technique used in TEL

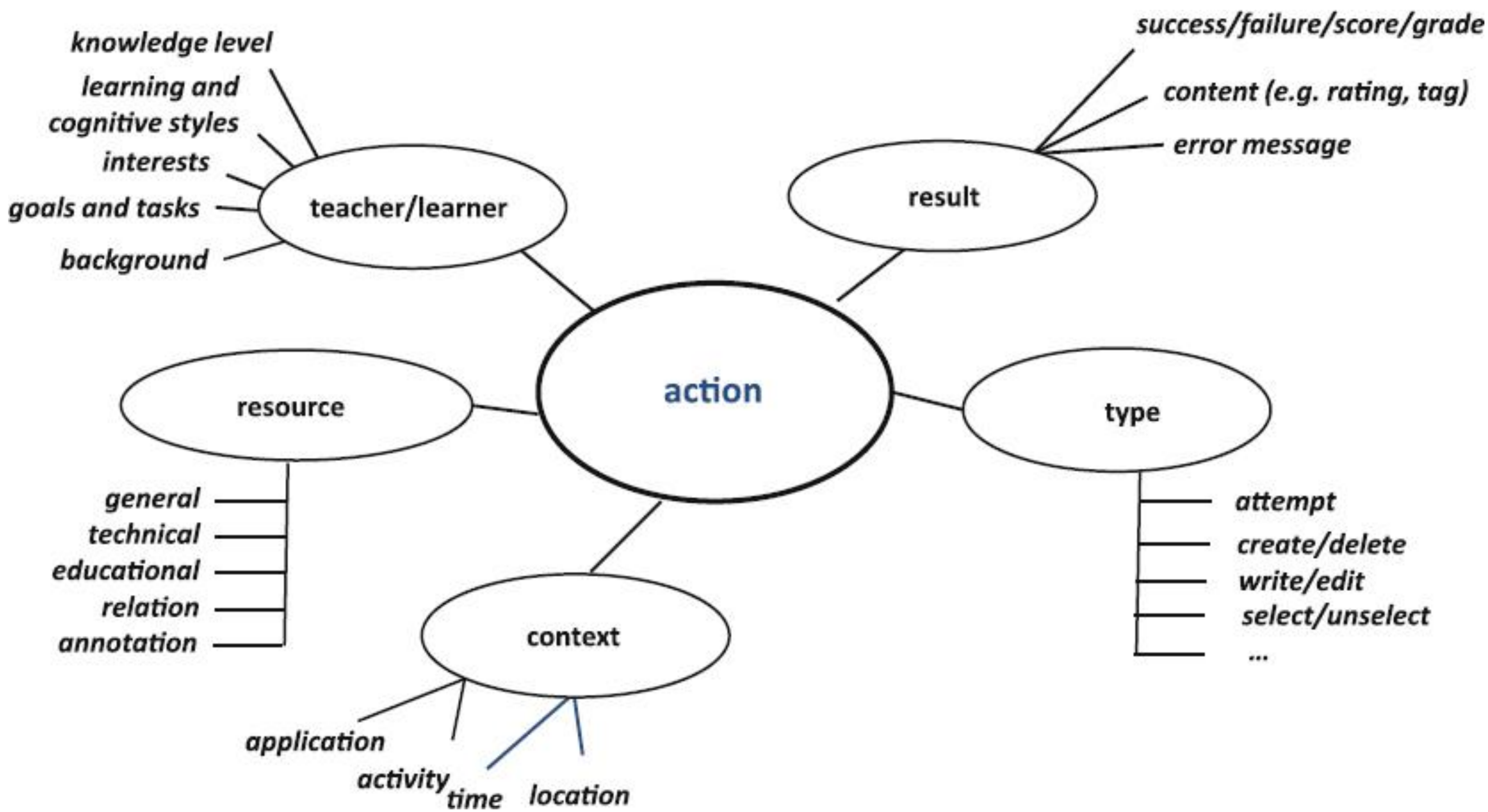
Content-based (CB) techniques				
Name	**Short description**	**Advantages**	**Disadvantages**	**Usefulness for TEL**
4. Case-based reasoning	Assumes that if a user likes a certain item, (s)he will probably also like similar items. Recommends new but similar items.	– No content analysis – Domain-independent – Quality improves over time	– New user problem – Overspecialization – Sparsity – Cold-start problem	– Keeps learner informed about learning goal – Useful for hybrid RS
5. Attribute-based techniques	Recommends items based on the matching of their attributes to the user profile. Attributes could be weighted for their importance to the user.	–No cold-start problem –No new user I new item problem –Sensitive to changes of preferences –Can include non-item related features –Can map from user needs to items	–Does not learn –Only works with categories –Ontology modeling and maintenance is required – Overspecialization	– Useful for hybrid RS – Recommendation from the beginning

TEL : Technology Enhanced Learning

Recommendation technique used in TEL

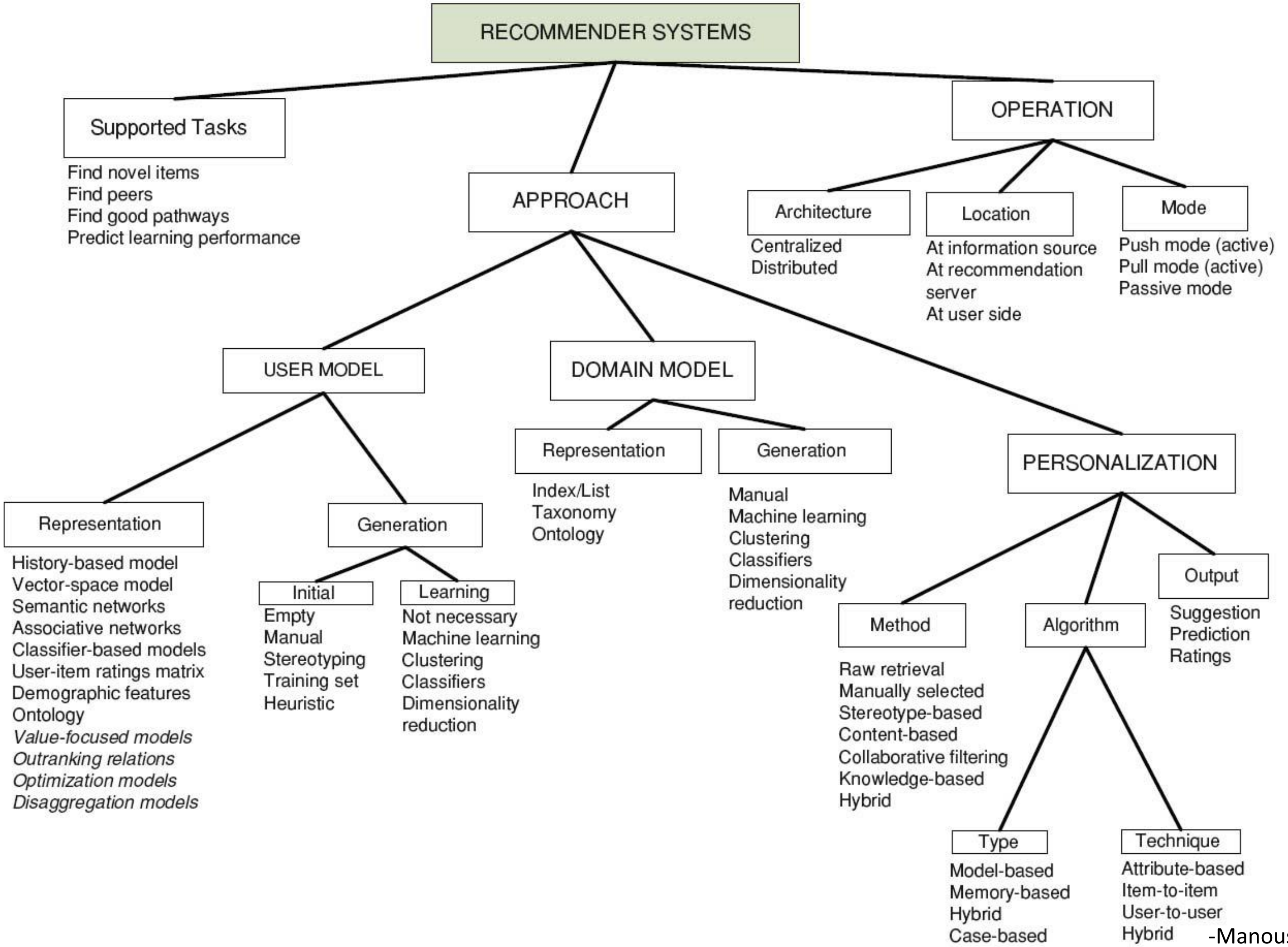
Data-Mining (DM) techniques				
6. Decision Trees (C4.5,ID3)	A decision tree represents a set of classifications created from a set of rules. They start form a single classification and branch out based on classification rules mined from the data.	–Easy to understand –High representation power	–Overspecialisation in small datasets –Can become very broad	–Visualize differences of learners from the data –Alternative approach to expert driven ontologies
7.K-Nearest Neighbor (Isodata, Forgy)	Does not build an explicit model instead exams the categories of the *K*-most similar data points. *K*-means is often used in TEL recommenders to compute similarity of vector-based approaches.	–Simple approach only two parameters to select –Robust to noise –High representation power	–Difficult to select distance function d –Irrelevant data needs to be removed –Slower than model-based recommendations	–Recommend similar peers, or contents to learners –Cluster learners in groups
8. Vector-based models (TF-IDF, Singular value decomposition, Matrix Factorisation)	Vector-based approaches characterise items and users as vectors of factors in a 3D space. A high correlation between an item and a user can be used as recommendation but also predictions can be created.	–Suitable for sparse datasets –Can take temporal differences into account –Can take various implicit information into account does not need explicit ratings	–Content depended (Items with same context but different terms are not matched) –User keywords have to match semantic space	–Useful to monitor and predict learner performance –Can adapt to increased knowledge level of learners –Can mark learning resources that are not popular anymore

TEL variables



- K. Verbert, N. Manouselis, H. Drachsler, E. Duval

Framework for the analysis of Recommender Systems



-Manouselis and Costopoulou (2007)

Prescribed Text Books

	Author(s), Title, Edition, Publishing House
	Recommender Systems for Learning by Nikos Manouselis, Hendrik Drachsler, Katrien Verbert and Erik Duval Springer © 2013
	Recommender Systems for Location-based Social Networks by Panagiotis Symeonidis, Dimitrios Ntempos and Yannis Manolopoulos Springer © 2014

BITS Pilani

Pilani | Dubai | Goa | Hyderabad

9.2 Fraud Detection

Source Courtesy: Some of the contents of this PPT are sourced from materials provided by publishers of prescribed books

Why study Fraud?

- Fraud can be defined as a criminal activity, involving false representations to gain an unjust advantage (*Concise Oxford Dictionary*)
- The Association of Certified Fraud Examiners estimates that U.S. organizations lose about 7% of their revenues to fraud. If this were to hold true for all organizations contributing to the Gross Domestic Product of about $14 trillion for 2007, fraud losses could be as high as $1 trillion

Issues with Fraud Detection

- Fraud is usually a rare event. Identifying fraud is difficult because of its rarity and because its very nature is stealthy.
- We need accurate models to make effective detection.
  - The vast majority of the records (i.e., 99.9%) may be legitimate. Only 0.1% of the records may be fraudulent. Here a 99% accurate model will lead to too many false alarms.
  - Say we have million transactions. As per above, 1000 are fraudulent. With 99% accuracy, i.e. 1% inaccuracy (false positives, false negatives), total alarms will be 1% of 0.999 million false alarms and 990 (99% of 1000) true alarms.

    Total alarms = 9990 + 990 = 10980, out of which more than 90% are false alarms.
- Often, the extra accuracy is associated with higher cost, but the cost of *not* doing so may be much higher

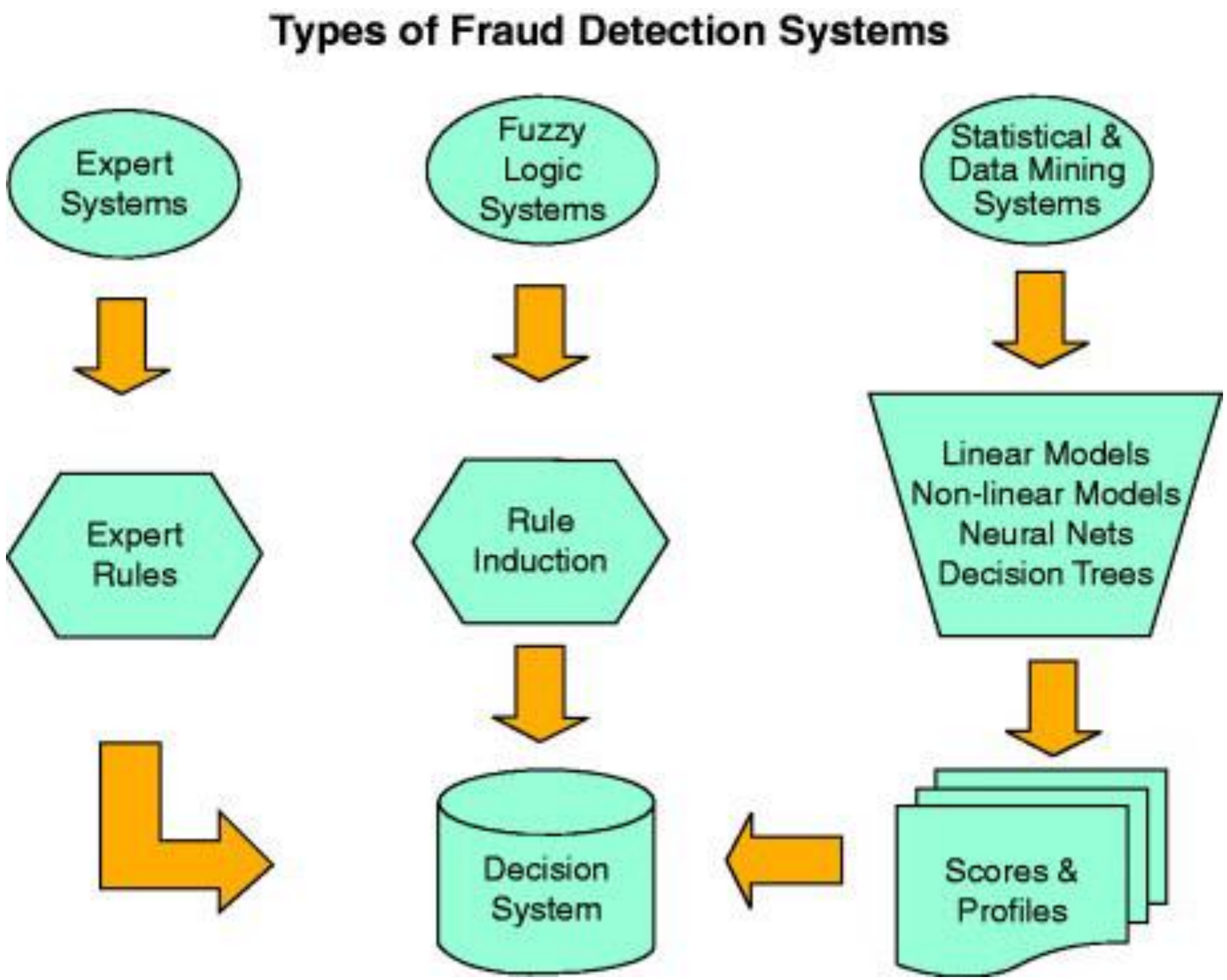
Issues with Fraud Detection

- **Fraud is Evolving**
  - Fraudsters may adapt quickly to many fraud detection methods, by devising novel and increasingly subtle ways to get away with it. Also, fraud detection schemes must evolve also to try to keep up with (and get ahead of) fraudsters
- **Large Data Set Processing Needed**
  - Large credit card issuers like Capital One may process billions of transactions per year. Even a very small percentage of fraud among these billions of transactions can result in proportionately large losses.
  - Telecom companies handle billions of calls in a month
- **The Fact of Fraud is Not Always Known during Modeling**
  - We need to use both supervised and unsupervised methods to detect fraud.
- **Fraud is Very Complex**
  - The complexity is partly due to the fraudster's need for stealth and secrecy, and partly due to the intentional obfuscation of the trail of evidence indicating fraud

Issues with Fraud Detection

- **Fraud Detection May Require the Formulation of Rules Based on General Principles, "Red Flags," Alerts, and Profiles**
  - **General principle:** The incidence of fraud is more likely when the opportunity is high and the potential gains are large.
  - **A "red flag":** A large number of accidents or claims is made by one individual
  - **An alert:** A new product is introduced before fraud management systems are put in place
- **Fraud Detection Requires Both Internal and External Business Data**
  - Internal data describing their business events (selling things or providing services)
  - External data such as demographic data, firmographic data (profile of businesses), psychographic data (people with various attitudinal and philosophical views)
- **Very Few Data Sets and Modeling Details are Available**
  - Fraud data sets and modeling methodologies are tightly kept secrets. Companies do not share with anyone.

Types of fraud models



Types of Fraud Models

- Early fraud models employed expert systems to detect fraudulent events. An expert system is a collection of expert opinions on a number of decision criteria. These systems induced rules from the responses of a group of experts in the field. These rules can be coordinated into a flow chart leading to a decision.
  - The problem with expert systems is that they are based on subjective inputs that may be contradictory

- Subsequent fraud detection systems used automated rule induction engines, based decision tree technology, and fuzzy logic. Some of these fraud detection systems are still marketed today (iPrevent by Brighterion).

Types of Fraud Models

- The Fair Isaac fraud detection systems Falcon Fraud Manager, eFalcon, and LiquidCredit Fraud Solution are built around a sophisticated system of predictive variables derived from extensive historical customer data.
  - These predictors have been selected by many years of modeling fraud in many companies. The variables are submitted to a powerful backpropagation neural net.

Fraud in Aspects of Business

- Fraud can occur in many aspects of business, for e.g.:
  - **Credit card fraud:** Stealing or counterfeiting credit card numbers, or nonpayment of accounts
  - **Application fraud:** Untrue statements on a credit application, leading to assignment of an artificially low credit risk
  - **Claim fraud:** Submitting inflated or false claims
  - **Life insurance:** False or "engineered" death claims
  - **Health care fraud:** False billings by health care providers

etc.

Supervised Methods for Fraud Detection

Several elements are crucial to the successful supervised fraud model

- The fraud event and the relationship of that event to specific transactions or responses of the fraudster must be accurately identified
- Historical data of past transactions or responses must be available to derive powerfully predictive variables
- Profiles of the past behavior and actions of both the fraudsters and the nonfraudsters must be built and employed in the modeling methodology

- Predictive variables need to be identified for each type of fraud.

Detection of money laundering and other financial crimes

To detect money laundering and other financial crimes, it is necessary to integrate information from multiple databases such as bank transaction databases, and federal or state crime history

Multiple data analysis tools can then be used :

- Data visualization tools to display transaction activities using graphs by time and by groups of customers
- Linkage analysis tools to identify links among different customers and activities
- Classification tools to filter unrelated attributes and rank the related ones
- Clustering tools to group different cases
- Outlier analysis tools to detect unusual amounts of fund transfers or other activities, and
- Sequential pattern analysis tools to characterize unusual access sequences

Identifying Tax Fraud through Social Network Analysis

- SNA(Social Network Analysis) is an analytic approach of correlating people, entities and relationships to determine how tightly an individual or business is related to others who have known compliance issues
- These relationships can be from shared
  - phone numbers,
  - physical addresses,
  - bank accounts,
  - credit cards, or
  - any other connection
- Graph analytics can give insights into members of the network

www.fico.com

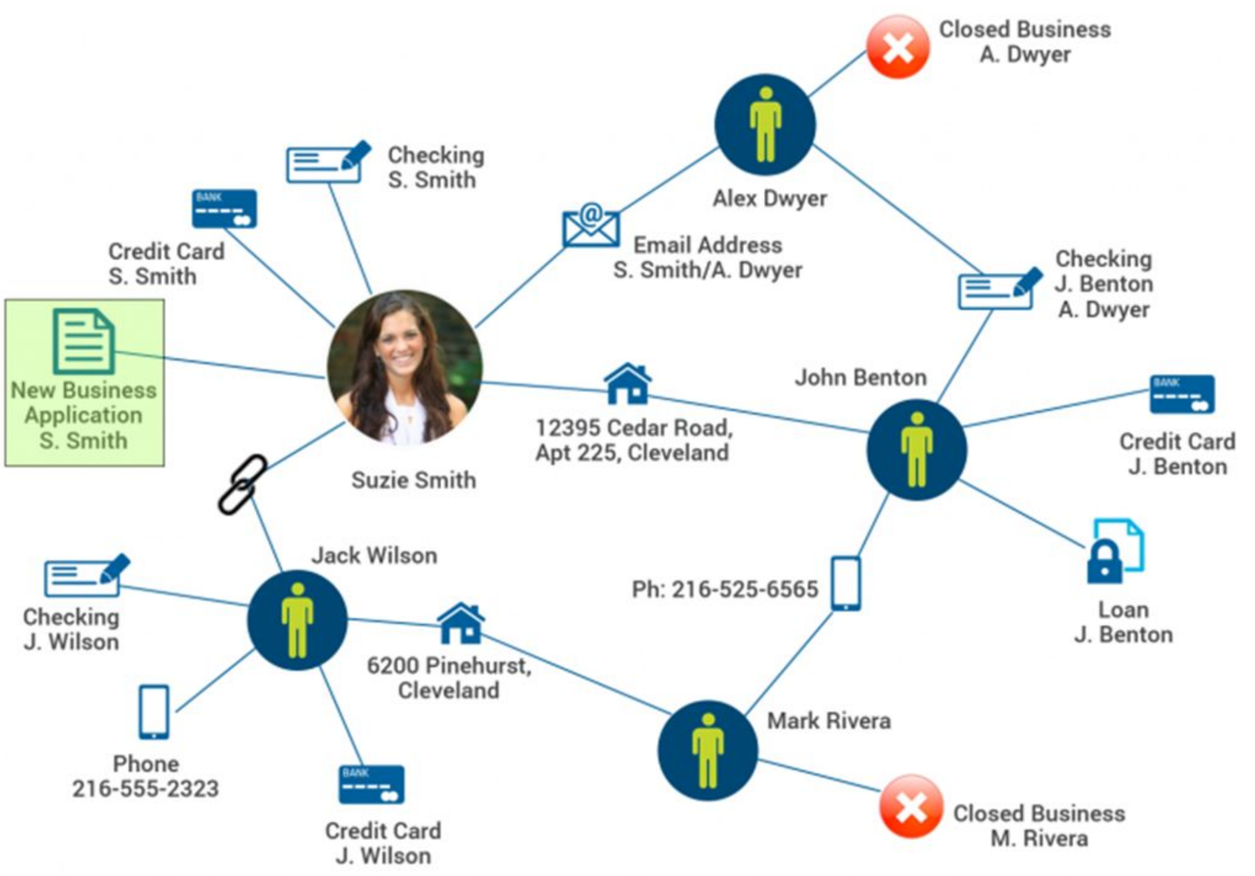
Opportunities for Tax Agencies in SNA

- Tax and revenue agencies to take advantage of SNA tools, across registration, audit and collections business areas

- **Improper registrations.** A tax evader closes a business and the business re-opens (typically owned by a relative of the original owner) at the same or a nearby location. i.e. the owner stays in business by opening a similar (or identical) business with a different legal name and a different legal owner (e.g., a spouse, parent or another relative).
  - Very difficult to catch by manual efforts. The challenge for the tax agency is the owner will have changed and won't be an exact match to the previous business.

www.fico.com

Opportunities for Tax Agencies in SNA



www.fico.com

- In the example, Suzie Smith has a direct relationship with one closed business and a second degree relationship with another closed business.
- The state can deny the new registration or conduct additional investigations in a cost effective manner.

Opportunities for Tax Agencies in SNA

- **Identify Fraud Rings.** The SNA can identify related businesses. i.e. businesses with related addresses, bank accounts, phone numbers, email addresses, or other identifying characteristics.
  - SNA can help a revenue agent can identify additional individuals and/or businesses that may be related to the same fraud ring, saving the investigator time and effort.
- **Assisting with Locating a Delinquent Debtor.** SNA's are frequently used during the debt collection process to identify related individuals. SNA can greatly enhance and automate this effort, by finding people who share the same physical address, phone number, email, etc
  - Earlier collectors have utilized manual tools along these lines for years, contacting next-door neighbors (for example) in an attempt to locate a debtor.

www.fico.com

Opportunities for Tax Agencies in SNA

- **Finding Successor Businesses.** When a business ceases operation, the business can re-open in a new location or under new ownership. If the original business owes money, the government can in many cases pursue that debt if there is a successor business
  - This can save the collector a significant amount of time for what otherwise would require significant manual research.

- Tax agencies are data rich organizations, and analytics solutions like Social Network Analysis will allow them to identify more fraud and potential non-compliance situations well before a liability occurs

www.fico.com

Prescribed Text Books

	Author(s), Title, Edition, Publishing House
	Fraud Detection; Handbook of Statistical Analysis and Data Mining Applications by Robert Nisbet, John Elder and Gary Miner Academic Press 2009
	Data Mining: Concepts and Techniques, Second Edition by Jiawei Han, Micheline Kamber and Jian Pei Morgan Kaufmann Publishers
	www.fico.com (Fair Issac Corporation)

innovate achieve lead

BITS Pilani

Pilani | Dubai | Goa | Hyderabad

9.3 Sentiment Analysis

Source Courtesy: Some of the contents of this PPT are sourced from materials provided by publishers of prescribed books

Sentiment Analysis

- Monitoring what consumers are saying about a company's brands and products and how they are expressing their opinions and sentiments to others has always been important to businesses.
- Until the last century, businesses typically used surveys and focus groups from time to time to gauge and track consumer sentiments.
- With the widespread adoption of the Internet, the proliferation of social media channels (such as Twitter, Facebook, and others), and the abundant opportunity for consumers to express their opinions and sentiments, monitoring sentiment continuously has become more critical
- "Conventional marketing wisdom long held that a dissatisfied customer tells ten people; but in the age of new social media, he or she has the tools to tell millions,"

  Paul Gillin, author of *The New Influencers: A Marketer's Guide to the New Social Media*

Sentiment Analysis

- The basic task involved in sentiment analysis is identifying and quantifying the polarity or valence of sentiments (such as positive, negative, neutral, or mixed) expressed typically in written opinions, expressions, reviews, comments, and so on
- It involves many of the text analytics steps such as
  - tokenization,
  - sentence identification,
  - part-of-speech tagging,
  - and so on

Sentiment Analysis

- Need to identify statements that convey sentiment

  "It is an amazing TV“ – conveys sentiment

  "Do not buy this TV" – conveys sentiment

  "Which is the best TV?" – conveys no sentiment

- Depending on the context, sentences can be non-comparative (where opinion is restricted to one thing) or comparative (where multiple things might be compared).

Example of a Review

Consider the following example of a review by a customer for a TV:

The **TV** is **wonderful**. **Great size**, **great picture**, **easy interface** It makes a **cute** little song when you **boot** it up and when you shut it off I just want to point out that the **43" does not** in fact play videos from the USB This is really **annoying** because that was one of the major perks I wanted from a new **TV** Looking at the product description now. I realize that the feature list applies to the **X758** series as a whole, and that each model's capabilities are listed below. Kind of a **dumb** oversight on my part, but it's equally **stupid** to put a description that does not apply on the listing for a very specific model

Granularity of Sentiment

Sentiment analysis starts with determining whether a text contains an opinion (sentiment). If it does contain sentiment, at what granularity level does the sentiment exist?

- *Document level:* At this level, the task is to figure out whether the entire document can be classified as positive or negative
  - This is possible only if the document involves a single entity (such as the TV in the previous example).
- *Sentence level:* At this level, the task is to classify each sentence in a document as a positive, negative, or mixed sentiment sentence.
  - In the previous example, the first sentence, expresses positive sentiment. The third statement, expresses negative sentiment

Granularity of Sentiment

Sentiment granularity can also be looked at from the object side:

- *Entity (or Object) and Attribute (or Aspect or Feature) level:* An entity is typically the target of the opinion. However, in many sentences, the sentiments reflect the reviewer's opinions about attributes (or aspects or features) of the entity
  - "Great size, great picture, easy interface", express positive sentiment for three specific attributes “size, picture, and interface” of the entity, the TV

Challenges in Sentiment Analysis - NLP

Sentiment analysis starts with text data. So it has all of the typical natural language processing (NLP) problems associated with text analytics, viz.

- identifying part-of-speech tags,
- disambiguating terms and lexicons,
- correcting spelling errors, etc.

In addition to words, there are idiom lexicons— e.g. "costs an arm and a leg" that embody sentiments.

In general, the difficulty with correctly identifying sentiments increases as you move from general to context-dependent to idiom lexicons in texts

Challenges in Sentiment Analysis – Opinion Words

Sentiment analysis needs to correctly identifying opinion words that express positive or negative sentiments.

- There are opinion words whose polarity is always the same, e.g. the word "beautiful," always expresses a positive sentiment.
- But, there are also context-dependent lexicons in which the polarity of the word depends on the domain or context, e.g. the word "small" can be positive or negative depending on the context. The sentence, "The size seems *small*." can be positive for a USB flash drive with 1 TB capacity. But, the same sentence can be interpreted as negative if the context is an LED big-screen TV.

Challenges in Sentiment Analysis – Type of Text

Further challenges in conducting sentiment analysis come from the nature of the text.

- For example, tweets are short, and they are typically focused on one topic only. In that sense, they are easier to analyze. But, tweets often contain a lot of special meaning characters, such as RT (retweets), hashtags (#), emoticons (such as smiley faces), that need to be handled carefully.
- Customer reviews are typically on one entity or object. Therefore, there is less ambiguity in the entity detection task when analyzing reviews.
- Analysis of discussions, free-flow comments, and blog postings is often the hardest because they typically cover multiple entities, make comparisons instead of expressing direct opinions, use a lot of sarcasm, etc.

Unsupervised versus Supervised Sentiment Analysis

- The sentiment analysis can be formulated as a supervised or an unsupervised mining problem, depending on whether there are known examples of documents belonging to positive or negative sentiments.
- Unsupervised sentiment analysis involves the application of a sentiment lexicon of opinion-related positive or negative terms to evaluate text in the document.
- Supervised approach involves machine-learning algorithms (such as support vector machines (SVMs) and neural networks) to textual feature representations to derive the relationships between features of the text segment and the opinions expressed in the document.
  - In many situations, known class examples are created by experts who read the documents or use rules, e.g. if a text review's numeric rating is four or more stars, then the review is positive.
  - If no known class examples are possible, then analysts have to use an unsupervised classification of sentiments.

Unsupervised versus Supervised Sentiment Analysis

Supervised classification is typically performed at the document level.

- If enough labeled examples are available, commonly used classification models can be trained, validated, and tested to check their performances.
- A good candidate is product review data, which typically has a text review and an overall numeric rating on a scale of one to five stars. Often, a review rating of four to five stars is considered a positive rating, and a review rating of one to two stars is considered a negative rating.
- The main challenge for modelers is to select the inputs from text features such as terms and their frequencies (often weighted or normalized), part-of-speech tags, opinion lexicons (general, context-specific, and idiom), syntactic dependency (from parsing trees), and the handling of negation words (such as "not").

Unsupervised versus Supervised Sentiment Analysis

Unsupervised method is typically applied at the sentence level. There are two types of unsupervised methods: lexicon-based and syntactic-pattern based.

- The *lexicon-based approach* can be used for sentence- and aspect-level sentiment classification. The relationships between opinion words and attributes are identified via dependency relationships obtained through parsing.
    - For example, in the sentence, "The picture quality is outstanding," the opinion word "outstanding" and the attribute "picture quality" share the same dependency relationship with the verb "is."
  - If a clear dependency is not observed between an opinion word and an attribute, then how close an opinion word is to an attribute in a sentence can be used to judge the polarity of the attribute.
    - This process can get very complex, depending on how long the sentence is, how many attributes are being mentioned in the same sentence, whether both positive and negative polarity words are used in the same sentence, whether negation is used, and so on.
  - Once sentiment values are computed for each word-attribute combination, they are typically combined using appropriate normalization or weights to come up with an overall sentiment score.
- The *syntactic pattern-based approach* involves defining part-of-speech tags and the keywords AND, NOR, OR, NOT, BUT, etc.
    - Primarily useful in contextual analysis when performing phrase -level analysis, this method can be used to develop a variety of rules for better accuracy. For example, a simple pattern such as <subject> <NOT> <verb> can be used to extract negative phrases like, *"This <feature> does <not> < work> as advertised."*

Prescribed Text Books

	Author(s), Title, Edition, Publishing House
	Tutorial BB - Mining Twitter for Airline Consumer Sentiment Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications by Gary Miner et al. Academic Press © 2012
	Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS SAS Institute © 201

innovate achieve lead

Thank You