Collaborative filtering: CF

Introduction : CF

- In everyday life, we rely on recommendations from other people either by word of mouth, recommendation letters, movie and book reviews printed in newspapers

- Recommender systems or recommendation systems are a subclass of information filtering system that seek to predict the 'rating' or 'preference' that user would give to an item (such as music, books, or movies) or social element (e.g. people or groups) they had not yet considered, using a model built from the characteristics of an item (content-based approaches) or the user's social environment (collaborative filtering approaches)

Contd…

- Recommender systems have become extremely common in recent years.

- Example :

  - When viewing a product on Amazon.com, the store will recommend additional items based on a matrix of what other shoppers bought along with the currently selected item.

Contd...

Recommender systems typically produce a list of recommendations in one of two ways –

1. Collaborative filtering approaches to build a model from a user's past behavior (items previously purchased or selected and/or numerical ratings given to those items) as well as similar decisions made by other users, then use that model to predict items (or ratings for items) that the user may have an interest in.

2. Content-based filtering approaches utilize a series of discrete characteristics of an item in order to recommend additional items with similar properties.

3. These approaches 1 and 2 are often combined (called as Hybrid Recommender Systems).

Contd…

- Recommender systems are a useful alternative to search algorithms since they help users discover items they might not have found by themselves.

- Interestingly enough, recommender systems are often implemented using search engines indexing non-traditional data.

- Recommender system is an active research area in the data mining and machine learning areas.

- They are costly to maintain
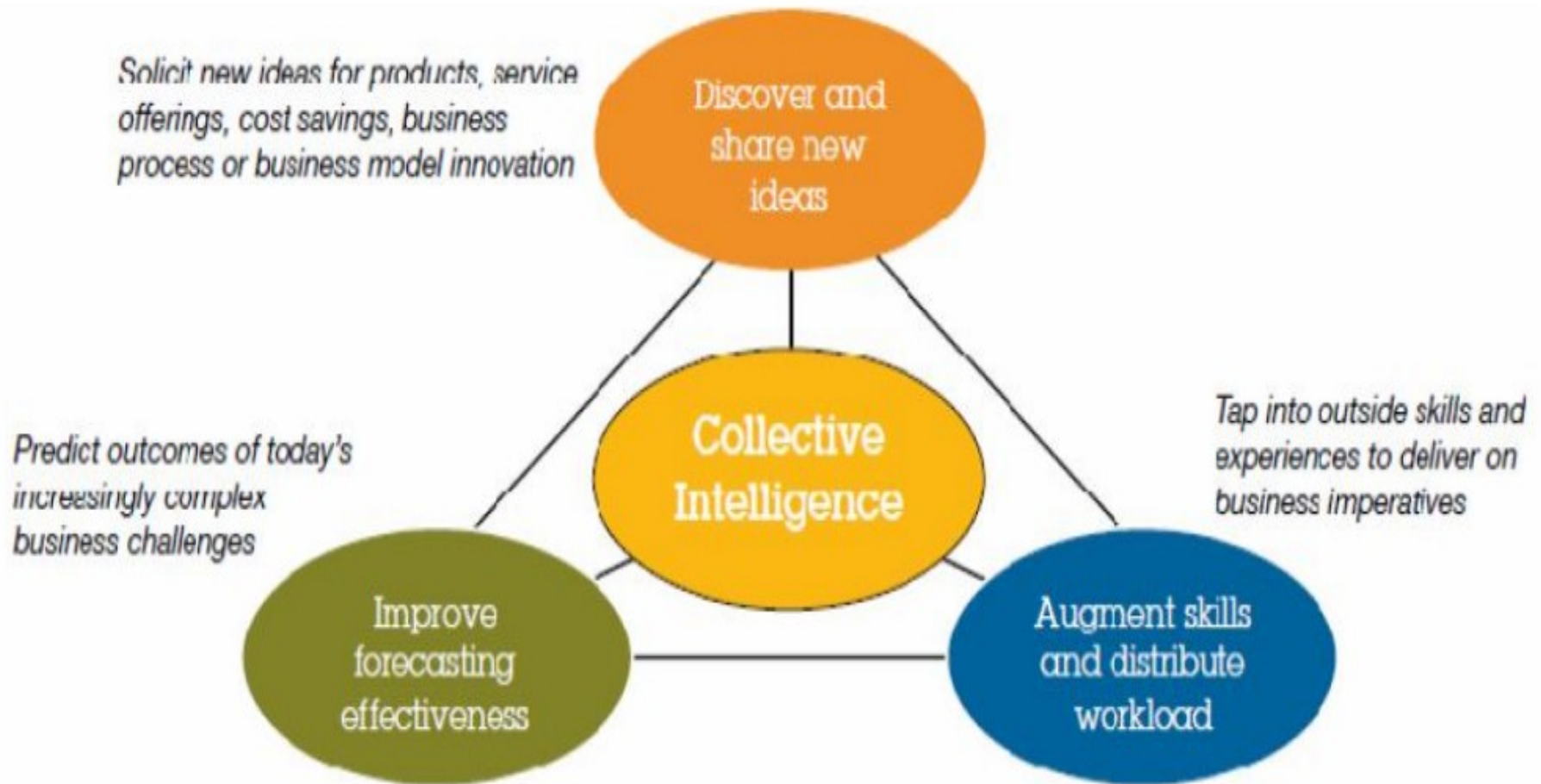
Motivation- Collective Intelligence

- Most complex systems, not only can be, but need to be viewed as collectives. Examples include:
  - Control of a constellation of communication satellites
  - Routing data/vehicles over a communication network/highway
  - Dynamic data migration over large distributed databases
  - Dynamic job scheduling across a (very) large computer grid
  - Coordination of rovers/submersibles on Mars/Europa
  - Control of the elements of an amorphous computer/telescope

Collective Intelligence

- Collective Intelligence – the aggregated knowledge, insight and expertise of a diverse group –has become a reality.
- As individuals become more adept and comfortable sharing thoughts and ideas in virtual spaces, companies can use these insights to address critical business challenges.
- Harnessing Collective Intelligence can play an important role in generating new ideas, solving age-old problems, disaggregating and distributing work in new and innovative ways, and making better, more informed decisions about the future.

Figure : Collective Intelligence

experience & knowledge to where it is most needed to address latest business

Solicit new ideas for products, service offerings, cost savings, business process or business model innovation

Discover and share new ideas

Tap into outside skills and experiences to deliver on business imperatives

Predict outcomes of today's increasingly complex business challenges

Collective Intelligence

Improve forecasting effectiveness

Augment skills and distribute workload

- COLLECTIVE
  - "Collective" refers to any entity constituted by other entities.
  - In this case, it usually refers to human social entities such as groups, organizations and communities.
- INTELLIGENCE AS A CAPACITY
  - Intelligence is variously defined as "the capacity to acquire and apply knowledge," "the ability to effectively adapt," or simply "the ability to solve problems."
  - Collective intelligence is an emergent property of collective social systems.
  - Collective intelligence is a natural product of the independent opinions or behaviors of diverse individuals or groups in a decentralized system (flock, market, guessing game) that aggregates those opinions or behaviors.

Contd...

- Collective intelligence is a theory that describes a type of shared or group intelligence that emerges from the collaboration and competition of many individuals and appears in consensus decision making in bacteria [clarification needed], animals, and computer networks [citation needed]

- The term appears in sociobiology, political science and in context of mass peer review and crowd sourcing applications.

- Everything from a political party to a public wiki can reasonably be described as this loose form of collective intelligence.

Contd…

- Primitive examples: families, companies, and countries etc.

- A precursor of the concept is found in entomologist –ant colony optimization.

- Collective intelligence is mass collaboration. In order for this concept to happen, four principles need to exist:

  – **Openness**: Sharing ideas and intellectual property: though these resources provide the edge over competitors more benefits accrue from allowing others to share ideas and gain significant improvement and scrutiny through collaboration

  – **Peering:** Horizontal organization as with the 'opening up' of the Linux program where users are free to modify and develop it provided that they make it available for others. Peering succeeds

Contd…

- **Sharing:**
  - Companies have started to share some ideas while maintaining some degree of control over others, like potential and critical patent rights.
  - Limiting all intellectual property shuts out opportunities, while sharing some expands markets and brings out products faster.
- Acting Globally
  - The advancement in communication technology has prompted the rise of global companies at low overhead costs.
  - The internet is widespread, therefore a globally integrated company has no geographical boundaries and may access new markets, ideas and technology.

Contd…

- In this context collective intelligence is often confused with shared knowledge. The former is knowledge that is generally available to all members of a community while the latter isinformation known by all members of a community.

 Examples:

- Google takes the collective knowledge created by millions of people making websites for other purposes and harnesses that collective knowledge—using some very clever algorithms and sophisticated technology—to produce amazingly intelligent answers to the questions we type in.

- Wikipedia, at another extreme, uses much less sophisticated technology, but some very clever organizational principles and motivational techniques, to get thousands of people all over the world to volunteer their time to create an amazing on-line collection of knowledge.

Contd...

- Social bookmarking:
  - In social bookmarking (also called collaborative tagging), users assign tags to resources shared with other users, which gives rise to a type of information organization that emerges from this crowd sourcing process.
  - Stock market predictions: Because of the Internet's ability to rapidly convey large amounts of information throughout the world, the use of collective intelligence to predict stock prices and stock price direction has become increasingly viable. Websites aggregate stock market information that is as current as possible so professional or amateur stock analysts can publish their viewpoints, enabling amateur investors to submit their financial opinions and create an aggregate opinion.
  - The opinion of all investor can be weighed equally so that a pivotal premise of the effective application of collective intelligence can be applied: the masses, including a broad spectrum of stock market expertise, can be utilized to more accurately predict the behavior of financial markets
  - on line gaming/ youtube etc...

Benefits of collective intelligence to business

- Talent Utilization
  - At the rate technology is changing, no firm can fully keep up in the innovations needed to compete.
  - Instead, smart firms are drawing on the power of mass collaboration to involve participation of the people they could not employ.
- Demand Creation
  - Firms can create a new market for complementary goods by engaging in open source community.
- Costs Reduction
  - Mass collaboration can help to reduce costs dramatically. Firms can release a specific software or product to be evaluated or debugged by online communities.
  - The results will be more personal, robust and error-free products created in a short amount of time and costs.

Contd…

The key question: How can people and computers be connected so that collectively they act more intelligently than any individual, group, or computer has ever done before?

- Programming Collective Intelligence takes you into the world of machine learning and statistics, and explains how to draw conclusions about user experience, marketing, personal tastes, and human behavior in general -- all from information that you and others collect every day. Each algorithm is described clearly and concisely with code that can immediately be used on your web site, blog, Wiki, or specialized application.

Contd…

- Popular methods:
  - Collaborative filtering techniques that enable online retailers to recommend products or
  - media
  - Methods of clustering to detect groups of similar items in a large dataset
  - Search engine features -- crawlers, indexers, query engines, and the PageRank algorithm etc.
- Understanding collective intelligence can help us fulfill the original dream of democracy: the participatory determination of our collective fate