Applications in Predictive Analytics:

Recommendation Systems

Northwestern University

Every day millions of people are subject to the power of persuasion. This persuasion is not rooted in any slick deals, clever advertising, or fast-talking salesmen, but in the power of their own choices. Enter: recommender systems – a technology that magically knows exactly what a user wants, very often before the user knows. Recommender systems, theorized in the 1970s and executed in the 1990s, can be defined as applications within computer software that possess the ability to make appropriate suggestions to users (Martin, Donaldson, Ashenfelter, Torrens & Hangartner, 2011, p. 19). Even those that get by with a minimal amount of technological savvy are aware of the success gained using the concept of recommendation by companies like Netflix, Amazon, and Pandora, though their techniques are not identical (Miller, 2015, p. 241). Currently, there exist three primary methods used by recommender systems: content-based systems, collaborative filtering, and a hybrid class (Taraghi, Grossegger, Ebner & Holzinger, 2013 p. 673). While all three of these methods cannot be applied to every organization or product, they certainly deserve equal consideration from any fascinated management due to their power to benefit the business, the user, or both.

Content-based recommender systems assign values to specific characteristics to items in order to determine a similarity index. For example, a company such as Netflix, specializing in movie rentals, may assign values to particular cast members, time/place settings, or movie categories. Likewise, Pandora may assign values to characteristics such as artist or genre. These values are often determined by authorities or specialists in the field (Taraghi et al., 2013, p. 674). Distinctive algorithms involving sparse matrices, that is, matrices comprised mostly of the zero value, are applied to this system, along with the user's personal characteristics, purchase history, actively given choices, and other defining facts, to generate for pertinent recommendations for the user (Miller, 2015, p. 242).

These systems have also been utilized as an alternative to standard search engines in fields such as academic research. Search engines are limited in their results from user queries as they are unable to determine in-depth content parallels along with direct text (Taraghi et al., 2013, p. 672). Content-based recommender systems, which do possess this ability, are an example of machine learning in action. Their distinctive algorithms continually absorb and process discernable patterns from data streams, producing a nearly instantaneous recommendation for even the newest of items (Henschen, 2013, p. 13). As a result, many recommender systems can focus on delivering unexpected results, actively promoting new, appropriate products, as well as facilitating an arena for the active information sharing in the educational community (Taraghi et al., 2013, p. 673).

The collaborative filtering method of recommendation operates on either a memory-based or model-based principle that the behavior of a crowd will help predict the behavior of an individual (Adomavicius & Tuzhilin, 2005, pp. 9 – 17). Algorithms in this setting, which are not content aware, search for consumers, or "neighbors," who have given similar ratings for an item to the user and recommend items that the neighbor has rated highly. In the same way, the algorithm may be used to search for other implicit data, such as similar product purchases to make a suggestion (Taraghi et al., 2013, p. 674). Data is not necessarily implicit, especially given the possible external influences or resources that can be identified with the user's interest. While proprietary data is the primary source, unrestricted public data, as well as accessible market research, should not be discounted with the rise of social media (Mela, 2011, p. 969). This type of recommender system has also been referred to as "social filtering," due to its community-dependent nature (Miller, 2015, p. 242).

The final method, a hybrid of the content-based and collaborative filtering recommendation systems, embrace what Dr. Eric Siegel, founder of Predictive Analytics World, refers to as “The Ensemble Effect,” which states, “When joined in ensemble, predictive models compensate for one another’s limitations, so the ensemble as a whole is more likely to predict correctly than its component models are” (Siegel, 2013, p. 149). Content-based systems are limited in that they simply bypass unique user personality, delivering recommendations that lack differentiation from items already used by the consumer. Collaborative filtering also sees its boundaries when neighboring users are unavailable; in addition, it fails to recommend new, possibly relevant items if those items are without a rating at the time. A hybrid recommendation system harnesses the best of both worlds by creating a combination model, being able to switch from one to the other, or by developing a flow or filter from a primary to secondary system of the organization’s choice, ultimately producing a more full-bodied suggestion (Taraghi et al., 2013, pp. 674-675).

Management should recognize that recommender systems are an extremely valuable application of predictive analytics. A 2012 survey of technology professionals by *InformationWeek* found that predicting customer behavior and predicting product sales topped the list of factors driving interest in big data analytics (Henschen, 2013, p. 8). These systems are beneficial both to users, helping them cope with information overload, and businesses, increasing revenues as technology costs steadily decline. With the future of recommender systems aimed at physiological awareness and cloud-based programs, most areas of management will not be disappointed with an investment in these methods (Martin et al., 2011 pp. 19 – 25). There are not many organizations that would turn down the opportunity to know users better than they know themselves, which is exactly what recommendation systems offer.

References

Adomavicius, G. & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge & Data Engineering*, *17*(6), 734-749. doi:10.1109/TKDE.2005.99.

Henschen, D. (2013). Analytics wish list. *InformationWeek,* (1361), 7-15. Retrieved from http://search.proquest.com/docview/1326779619?accountid=12861.

Martin, F. J., Donaldson, J., Ashenfelter, A., Torrens, M., & Hangartner, R. (2011). The big promise of recommender systems. *AI Magazine, 32*(3), 19-27. Retrieved from http://search.proquest.com/docview/908412542?accountid=12861.

Mela, C. F. (2011) Structural workshop paper – data selection and procurement. *Marketing Science 30*(6), 965-976. http:// dx.doi.org/10.1287/mksc.1110.0650.

Miller, T. W. (2015). *Modeling techniques in predictive analytics: Business problems and solutions with R* (revised and expanded ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Siegel, E. (2013). *Predictive analytics: The power to predict who will click, buy, lie, or die*. Hoboken, NJ: John Wiley & Sons, Inc.

Taraghi, B., Grossegger, M., Ebner, M., & Holzinger, A. (2013). Web analytics of user path tracing and a novel algorithm for generating recommendations in open journal systems. *Online Information Review, 37*(5), 672-691. doi:http://dx.doi.org/10.1108/OIR-09-2012-0152.