

۱ مقدمه

در عصر جدید اینترنت جایگاه ویژه ای در زندگی افراد پیدا کرده است. بسیاری از افراد کاملاً به اینترنت وابسته شده و کارهای متعددی از امور روزمره خود مانند خرید اقلام مورد نیاز، تراکنشات مالی، سرگرمی و... را از طریق اینترنت انجام می دهند. یکی از موارد بسیار مهم در اینترنت تعامل بین مشتری و فروشنده است. امروزه فروشندگان برای باقی ماندن در بورس رقابت برای فروش محصول خود به مشتریان باید امکان خرید اینترنتی کالاها را فراهم نمایند. مشتریان به راحتی در دنیای تجارت الکترونیک می توانند به هر فروشگاه و فروشنده ای که بخواهند سر بزنند و این موضوع رقابت بین فروشندگان را بسیار زیاد کرده است. برای کسب رضایت مشتری امروزه امکانات لازم برای خرید اینترنتی تنها یک مورد لازم می باشد و کافی نیست بلکه برای اینکار باید با ایجاد رابطه خوب با مشتری امکان خریدهای بعدی وی را بیشتر نماییم. با ایجاد رابطه مناسب با مشتری می توانیم امکان وفاداری مشتری را ارتقا دهیم که این مساله برای سازمان از اهمیت بالایی برخوردار است. از جمله کارهایی که به منظور افزایش رضایت مشتری و ایجاد وفاداری در وی انجام می شود شخصی سازی است. شخصی سازی مفهومی در مدیریت ارتباط با مشتری است که هدفش ایجاد یک رابطه مستقیم و رفتار خاص سازمان با یک مشتری خاص است که موجب دوام بیشتر وفاداری مشتری می شود. یکی از نموده های بارز شخصی سازی در دنیای تجارت الکترونیکی سیستم های پیشنهاد دهنده می باشند.

سیستم های پیشنهاد دهنده به دنبال ارائه محصول مناسب با توجه به خصوصیات و رفتارهای مشتری هستند. در صورتی که این مساله به خوبی انجام شود و به هر مشتری محصول و سرویس متناسب با خودش ارائه شود و به نوعی شخصی سازی به خوبی انجام شود، رضایت مشتری افزایش یافته که برای سازمان بسیار مهم می باشد. مساله بسیار مهم در این زمینه دقت روش پیش بینی کننده رفتار مشتری است. زیرا ارائه مناسب محصول و به نوعی پیشنهاد محصول مناسب و باب میل مشتری تاثیر عمیقی در دید و رضایت مشتری دارد ولی پیشنهاد بد و نامناسب به مشتری می تواند موجب نارضایتی و سرد شدن رابطه وی با سازمان شود.

برای انجام مناسب عملیات سیستم های پیشنهاد دهنده از جمله ابزارهایی که به کار گرفته شده است، امکاناتی است که داده کاوی در اختیار قرار می دهد. داده کاوی تکنیک های بسیار قوی برای کشف الگوهای کشف نشده در محصولات و مشتریان در اختیار سیستم های پیشنهاد دهنده قرار می دهد. از جمله تکنیک

هایی که داده کاوی معرفی می نماید کلاس بندی^۱، خوشه بندی^۲، قوانین همبستگی^۳، پیش بینی^۴ و... می باشد.

در این زمینه تحقیقات بسیار زیادی انجام شده است و افراد بسیار زیادی برای پیش بینی نیاز و رفتار مشتری از تکنیک های مختلفی استفاده کرده اند، اما این مساله همچنان باز است و تلاش برای بهبود روش های موجود و معرفی روش های جدید تر همچنان ادامه دارد. به طور کلی برای پیش بینی نیاز رفتار مشتری در سیستم های پیشنهاددهنده از چهار روش Collaborative filtering (CF) و (CBF) Hybrid Approach contend-based filtering و Demographic-based Method استفاده شده است. به طور کلی روش CF بر اساس اطلاعات بدست آمده از سایر مشتریان برای اظهار نظر در مورد مشتری جاری استفاده می کند و روش CBF از اطلاعات مربوط به رفتارها و خریدهای قبلی خود مشتری جاری استفاده می کند، Demographic-based Method از اطلاعات دموگرافیک کاربر مثل سن و جنس و شغل وی استفاده کرده و در روش Hybrid Approach نیز از جفت آنها استفاده می شود.

۲ تعریف مساله

مساله اساسی در سیستم های پیشنهاد دهنده این است که می خواهیم نظر یک مشتری را برای محصولی که تاکنون اظهار نظری در مورد آن نکرده است را پیش بینی نماییم.

به طور کلی صورت این مساله به شکل زیر می باشد:

یک مجموعه M شامل کالاها داریم به همراه یک مجموعه U شامل کاربران و یک مجموعه T شامل امتیازات:

$$M = \{m_1, m_2, m_3, m_4, \dots\}$$

$$U = \{u_1, u_2, u_3, u_4, \dots\}$$

$$T = \{t_1, t_2, t_3, t_4, \dots\}$$

برای هر کاربر یک مجموعه دو تایی به صورت (m, t) وجود دارد که می گوید کاربر جاری به کالای m امتیازی برابر t داده است:

$$\forall u_i : \{(m_1, t_1), (m_2, t_2), \dots\}, \quad u_i \in U, m_i \in M, t_i \in T$$

classification^۱
Clustering^۲
Association Rules^۳
Prediction^۴

مساله ای که سیستم های پیشنهاد دهنده به دنبال آن هستند این است که علاقه مشتری را با استفاده از این اطلاعات موجود پیش بینی نمایند و به نوعی علاقه مشتری را در مورد یک محصول خاص حدس بزنند تا در صورتی که با توجه به پیش بینی انجام شده حدس می زنیم این مشتری محصول مربوطه را می پسندد به وی پیشنهاد خرید و یا مشاهده آن محصول را ارائه دهیم. این کار اگر به درستی انجام شود باعث می شود بر اساس نظریات موجود در مبحث مدیریت روابط با مشتری رابطه مناسب بین مشتری و سازمان می شود و این امر باعث ایجاد وفاداری بیشتر و اعتماد بیشتر مشتری به سازمان می شود. این امر به طور طولانی مدت منجر به جلوگیری از ترک فروشنده توسط مشتری و سود بیشتر فروشنده می شود. اما به همان اندازه که انجام درست این کار خوب و مفید است انجام دادن نامناسب آن نیز می تواند رابطه بین مشتری و فروشنده را بد کرده و در دراز مدت سود فروشنده را کاهش دهد.

به طور کلی مساله اساسی در سیستمهای پیشنهاد دهنده پیش بینی علاقه مشتری به کالایی است که تا به حال نظر خود را در مورد آن اعلام نکرده است.

۳ مرور ادبیات

۳,۱ Collaborative filtering

CF یکی از مهم ترین و موفق ترین روش های استفاده شده در سیستم های پیشنهاددهنده می باشد. تحقیقات زیادی در این زمینه انجام شده است. به طور کلی در این روش با استفاده از امتیازاتی که سایر مشتریان به یک کالای خاص داده اند و تشابه بین مشتری جاری و سایر مشتریان سعی می شود امتیازی که مشتری به این کالای خاص می دهد پیش بینی شود. [1,2] GroupLens system و [3] Amazon.com از جمله نمونه های موفق از این نوع سیستم های پیشنهاددهنده می باشند.

برای پیش بینی امتیاز مشتری جاری باید مشتریانی که تشابه خوبی با وی دارند و امتیازات آنها را داریم شناسایی شده و بر اساس آنها نظر دهیم. برای تشابه مشتریان از معیارهای زیادی استفاده شده است. معیارهای مهمی که در کارهای قبل از [4] استفاده می شده است ، معیار Pearson's Correlation و Cosine و Adjusted Cosine for similarity between items و Constrained Pearson's Correlation و Spearman's Rank Correlation بوده است. **برای این درس شما باید فاصله اقلیدسی را بتوانید پیدا نمایید.**

برای محاسبه این فاصله تنها امتیازهای هر مشتری به کالاها مورد نظر می باشد در واقع فقط امتیازی که کاربر مربوطه به یک سری کالای های استفاده شده داده است با امتیاز کاربران دیگر به همان کالاها مقایسه می شود. و این فاصله به بقیه اطلاعات موجود کاری ندارد. وقتی رکورد جدیدی که شامل یک کاربر خاص و یک کالای خاص است که تا به حال علاقه خود را در مورد آن اظهار نکرده است را در نظر بگیریم ، اینکه چه افرادی کمترین فاصله را در اظهار علاقه به کالاهای قبلی به این مشتری دارند را یافته و بر اساس میانگین علاقه آنها به این کالای مورد نظر نظر خود را در مورد این ورودی اعلام می کنیم. **ممکن است روشهای دیگری مثل میانگین وزن دار نیز مورد استفاده قرار گیرد. شما باید هر دو را بتوانید انجام دهید.**

یکی از مشکلاتی که در روش CF وجود دارد مشکل cold-start می باشد. این مشکل موقعی پیش می آید که برای مشتری اطلاعاتی در مورد رفتار گذشته در اختیار نیست و این امر موجب می شود تا این روش نتواند مقایسه ای بین مشتری جدید و مشتریان سابق انجام دهد و در نتیجه در ارائه پاسخ باز می ماند همچنین زمانی که اطلاعات موجود از رفتار مشتری کافی نباشد و این مشتری تنها تعداد معدودی با فروشنده تعامل داشته است نیز این مشکل با شدت کمتری رخ می نماید، در این شرایط پیش بینی سیستم با مشکل دقت پایین مواجه خواهد شد.

مشکل دیگری که در این زمینه می توان به آن اشاره کرد این است که اگر اطلاعات مشتریانی که در اختیار داریم خیلی پراکنده باشند (مشتریان اشتراک زیادی در محصولات که امتیازدهی کرده اند نداشته باشند و هر کاربری اقلامی خاص را که افراد کمی آنها را امتیاز دهی کرده اند امتیازدهی کرده باشد) در چنین مواقعی الگوریتم اطلاعات کافی برای مقایسه امتیازهای فرد جاری و مشتریان قبلی ندارد و نمی تواند آنطور که باید نتیجه خوبی ارائه دهد. در چنین مواقعی بهتر است از روش CBF استفاده کرد.

۳.۲ contend-based filtering

در این روش از اطلاعات خرید ها و رفتارهای سابق خود مشتری جاری برای پیش بینی رفتار ش استفاده میشود . به این صورت که مثلاً موقع پیش بینی امتیازی به کالای خاصی که تا به حال مشتری جاری امتیاز نداده است از امتیازات سابق وی به کالاهای قبلی ای که خریده است استفاده می شود. به این صورت که کالاهای مشابه با کالای جاری شناسایی شده و بر اساس امتیازی که فرد قبلاً به آنها داده است برای امتیاز به کالای جاری استفاده می شود. روش های مختلفی برای این منظور به کار گرفته شده است. از جمله روشهای مهم استفاده شده می توان به الگوریتم knn ، روش های خوشه بندی ، شبکه عصبی و Association rule اشاره نمود. **برای یافتن کالاهای مشابه شما باید بتوانید بر اساس فاصله اقلیدسی اینکار را انجام دهید.**

مشکل cold-start در این روش نیز وجود دارد و اگر اطلاعات در مورد رفتار قبلی مشتری نداشته باشیم عملاً قادر به اعلام نتیجه ای در مورد پیش بینی رفتار مشتری ارائه نماییم. همچنین اگر این اطلاعات کم باشند و تعداد معدودی خریده‌ها و رفتارهای قبلی از مشتری ثبت شده باشد با مشکل دقت پایین الگوریتم مواجه خواهیم شد.

۳, ۳ Demographic-based Method

این روش مشابه روش CF سعی بر آن دارد تا از امتیازدهی سایر کاربران برای پیش بینی امتیاز کاربر جاری به کالای مربوطه استفاده نماید. فرق آن این است که در این روش شباهت بین افراد بر اساس اطلاعات دموگرافیک محاسبه می شود نه امتیازهای داده شده به کالاهای مشابه.

۳, ۴ Hybrid Approach

در این روش سعی بر آن است تا از مزیت های روشهای قبلی استفاده شود و نتیجه بهتری ارائه گردد. به طور کلی می توان سه نوع از روش ترکیبی نام برد. در نوع اول سعی می شود تا یکی از دو روش CF, CBF استفاده شود و قسمتی از اجزای روش دوم را ایجاد نماید مثلاً در مقاله [15] ابتدا از روش CBF استفاده شده است تا امتیازهای اضافی برای روش CF ایجاد نماید و سپس از روش CF برای پیش بینی امتیاز نهایی استفاده کرده است. در نوع دوم خروجی روش های دیگر بعد از انجام جداگانه هر کدام ترکیب می شوند تا نتیجه نهایی شکل گیرد. مثلاً در مقاله [16] نتیجه دو روش ابتدا بدست آمده و نهایتاً نتیجه نهایی از ترکیب خطی هر دوی آنها بدست آمده است. در نوع سوم یک مدل جامع بر اساس سایر روشها ایجاد می شود.

۴ منابع

- [1] J.A. Konstan, B.N. Miller, D. Maltz, J.L. Herlocker, L.R. Gordon, J. Riedl, GroupLens: applying collaborative filtering to use news, Communications of the ACM 40 (1997) 77–87.
- [2] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl, GroupLens: an open architecture for collaborative filtering of netnews, in: Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work. ACM, Chapel Hill, NC, 1994, pp. 175–186.
- [3] G. Linden, B. Smith, J. York, Amazon.com recommendations: item-to-item collaborative filtering, IEEE Internet Computing 7 (2003) 76–80.
- [4] Hyung Jun Ahn, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, Information Sciences 178 (2008) 37–51
- [5] Resnick P, Iacovou N, et al, "GroupLens: an open architecture for collaborative filtering of netnews", Proceedings of the ACM Conference on Computer Supported Cooperative Work, pp175-86, 1994.
- [6] Cho, Y.H., J.K. Kim, et al, "A personalized recommender system based on web usage mining and

decision tree induction", Expert Systems with Applications.Vol.23, No. 3, pp329-342, 2002.

[7] Aggarwal, C.C., J.L. Wolf, et al, "Horting hatches an egg: a new graph-theoretic approach to collaborative filtering", Proceeding of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp201-212, 1999.

[8] Min, S.H., I. Han, "Recommender systems using support vector machines", Proceeding of International Conference on Web Engineering, pp387-93, 2005.

[9] Breese, J., D. Heckerman, et al, "Empirical analysis of predictive algorithms for collaborative filtering", Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, pp43-52, 1998.

[10] D.H. Kim, V. Atluri, et al, "A click stream-based collaborative filtering personalization model: towards a better performance", Proceedings of the 6th Annual ACM International Workshop on Web Information and Data Management, pp88-95, 2004.

[11] Balabanovic, M., Y.Shoham, "Fab: content-based, collaborative recommendation", Communications of the ACM, Vol. 40, No. 3, pp66-72, 1997.

[12] Xu, B., M. Zhang, et al, "Content-based recommendation in E-Commerce", Proceeding of ICCSA, pp946-55, 2005.

[13] Mooney, R.J., L. Roy, "Content-based book recommending using learning for text categorization", Proceeding of ACM SIGIR Workshop Recommender Systems: Algorithms and Evaluation, pp195-204, 1999.

[14] Zhang, Y., J. Callan, et al, "Novelty and redundancy detection in adaptive filtering", Proceeding of ACM SIGIR, pp.81-88, 2002.

[15] Melville, P., R.J. Mooney, et al, "Content-boosted collaborative filtering for improved recommendations", Proceeding of Conference on Artificial Intelligence, pp187-92, Edmonton, 2002.

[16] Claypool, M., et al, "Combining content-based and collaborative filters in an online newspaper", Proceeding of ACM Workshop Recommender Systems: Algorithms and Evaluation, 1999.

[17] Popescul A, L.H. Ungar, et al, "Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments", Proceeding of 17th Conference on Uncertainty in Artificial Intelligence, pp437-44, 2001.

[18] Jong-Seok Lee, Chi-Hyuck Jun, Jaewook Lee, Sooyoung Kim, Classification-based collaborative filtering using market basket data, Expert Systems with Applications 29 (2005) 700–704

[19] <http://www.grouplens.org/>