

تجارت الکترونیکی تلفیقی و فرآیند تعیین الگوهای سودمند تجاری (Data Mining) :

معماری و چالش ها

(سهیل انصاری ، ران کوهاری ، لیلو می سن و زیجیان زنگ )

شرکت نرم افزارهای Blue Martini

چکیده :

ما نشان می دهیم که حوزه تجارت الکترونیکی می تواند همه اجزا مورد نیاز برای یک

فرآیند موفق تعیین الگوهای سودمند تجاری را فراهم سازد و در عین حال عقیده داریم که

این یک حوزه موثر برای فرآیند ذکر شده است . ما بر اساس تجاربمان در شرکت نرم

افزاری بلو مارتینی ( Blue Martini ) در ارتباط با معماری تلفیقی توضیح خواهیم داد

تا نشان دهیم که این امر بر اساس تجربه صورت گرفته است .

معماری به طور قابل توجهی می تواند زمان انجام پیش پردازش ، پاک سازی و اقدامات

شناخت داده ها را کاهش دهد که معمولاً این کارها حدود ۸۰٪ زمان پروژه های کشف

اطلاعات را در بر می گیرد . به منظور پشتیبانی از ثبت داده ها و فوق داده ها که برای

فرآیند کشف مفید است ، ما تاکید می کنیم که جمع آوری داده ها در لایه سرویس

دهنده کاربردی صورت بگیرد نه در سرویس دهنده وب.

توضیح داده خواهد شد که پل های انتقال داده بین سیستم های پردازش گردش کار یا جریانهای رخدادی خریدار ( به طور مثال : جریان کلیک) و محل انبار داده ها مورد نیاز است همچنین به تفصیل در ارتباط با استخراج اطلاعات ، کاری که نیاز مند داشتن چندین دیدگاه از داده ها از طریق گزارش، الگوریتم های data mining تجسم و OLAP می باشد ، توضیح خواهیم داد . و با بیان چندین چالش در این زمینه مقاله خود را به پایان خواهیم برد .

OLAP : یک سیستم بانک اطلاعاتی رابطه ای ، با قابلیت پردازش پرس و جوهای پیچیده تری نسبت به بانکهای رابطه ای استاندارد که از طریق دسترسی چند بعدی به داده ها ( که در آن داده ها با معیارهای مختلفی مورد توجه قرار می گیرند) قابلیت محاسبه زیاد و روشهای ویژه اندیس گذاری انجام می شود .

مقدمه :

تجارت الکترونیکی در حال پیشرفت است و همگام با این پیشرفت ، شرکت ها علاقه بیشتری دارند تا در زمینه پیشرفت تجارت آن لاین (ONLINE) سرمایه گذاری کنند .

در کتاب COMMERCE SOFTWARE TAKESOFF ، نویسندگان تحقیق

forrester نوشتند که در سال ۱۹۹۹ تجارت آن لاین در زمینه خرده فروشی مصرف

کننده حدود ۲۰/۳ میلیارد دلار بوده است و پیش بینی شده است که تا سال ۲۰۰۳ به بیش از ۱۴۴ میلیارد دلار برسد. حدود ۲۵۰۰ شرکت جهانی در سال ۲۰۰۰ میلادی نسبت به سال ۱۹۹۹ بیش از ۷۲٪ در زمینه تجارت الکترونیکی سرمایه گذاری خواهند کرد. سایت های موجود از اندازه گیرهای اولیه همچون تعداد بازدید صفحات استفاده می کنند. اما نیاز برای تحلیل و خصوصی سازی با توجه به نیازهای متفاوت روز به روز افزایش می یابد. در کتاب Measuring websuccess نویسندگان ادعا می کنند که افراد پیشرو از واحدهای متدی برای ایجاد خصوصی سازی استفاده می کنند و اینکه شرکت ها به هوش شبکه نیاز دارند نه به تحلیل موارد ثبت شده. ابزار data mining به کشف الگوهای داده ها کمک می کنند. تاکنون شرکتها یی که بر روی ساختار افقی وسایل الگو ساز data mining متمرکز شده بودند، پیشرفت تجارت کمی داشتند. میزان خرید بسیاری از شرکت ها بدین قرار است: دریافت علوم فشرده توسط Gentia در حدود ۳ میلیارد دلار، هماهنگی بسیار بالا توسط yahoo در حدود ۲/۳ میلیون دلار، معتدل سازی به وسیله spss در حدود ۷ میلیون دلار و ماشین های متفکر بوسیله oracle در حدود چیزی کمتر از ۲۵ میلیون دلار افزایش یافته اند. KD1 توسط Net Perception در حدود ۱۱۶ میلیون دلار به دست آمد. right point که

قبلاً نام data mind داشت توسط E.Piphany برای حدود ۴۰۰ میلیون دلار و Nero vista توسط Acro برای حدود ۱۴۰ میلیون دلار به دست آمده است. تغییر این ارزیابی‌ها تاکید می‌کند که باید توجه بیشتری را بر ارزش تکنیک‌های الگو سازی data mining در زمینه تجارت الکترونیکی داشته باشیم.

تجارت الکترونیکی یک حوزه موثر برای data mining است. این روش ایده‌آل است زیرا بسیاری از مواد مورد نیاز برای data mining موفق به راحتی به دست می‌آید: ثبت داده‌ها بسیار متعدد است، مجموعه الکترونیک داده‌های قابل اطمینان را فراهم سازد، افکار به راحتی به مرحله اجرا در می‌آید و بازگشت سرمایه اندازه‌گیری می‌شود. اگر چه برای پیشرفت در این حوزه، باید data mining در سیستم‌های تجارت الکترونیکی با پل‌های انتقال داده‌ها از سیستم پردازش گردش کار به محل انبار داده‌ها و بر عکس مجتمع شود. چنین ترکیبی به طور قابل توجهی زمان آماده‌سازی داده‌ها را کاهش می‌دهد که طی تحقیقات بدست آمده حدود ۸۰ درصد زمان تکمیل یک تحلیل را بر عهده می‌گیرد. یک راه حل تلفیقی همچنین می‌تواند برای کاربران یک رابطه مستقیم بین کاربر و پردازش کامل فوق داده‌ها را ایجاد کند.

این تحقیق بر اساس مواردی که بعداً ذکر خواهد شد، شکل گرفته است. بخش ۲ معماری تلفیقی را توصیف می کند و در مورد اجزای سازنده مهم و پل هایی که آنها را به هم وصل می کنند توضیح می دهد. بخش ۳ به تفصیل در مورد جمع آوری داده ها صحبت می کند. یعنی باید داده های بیشتری را نسبت به فایل های ثابتی سرویس دهنده شبکه تاکنون مورد استفاده بوده اند، جمع آوری کرد. بخش ۴ در مورد اجزای تحلیل گر صحبت می کند که باید طیف وسیعی از وسایل انتقال داده ها و وسایل آنالیز کردن را فراهم کرد. و در ارتباط با بعضی از مشکلات چالش انگیز در بخش ۵ صحبت خواهیم کرد. و در بخش ۶ با یک جمع بندی مطلب را به پایان خواهیم برد.

### معماری تلفیقی :

در این بخش ما در مورد معماری مورد نظر یک بازبینی کلی در زمینه سیستم تجارت الکترونیکی با استفاده از data mining تلفیقی خواهیم داشت. از جزئیات بخش های بسیار مهم معماری و سودمندی آنها در بخش های بعدی بحث خواهیم کرد. این سیستم که در مورد آن شرح دادیم یک معماری ایده آل است که براساس تجارب ما در شرکت نرم افزاری Blue martini بدست آمده است. اگر چه ما هیچ گونه ادعایی نداریم که هر آنچه در اینجا توضیح داده می شود آنرا در شرکت خود اجرا کرده ایم. در معماری مورد

نظر ما سه جزء بسیار مهم وجود دارد: تعریف داده تجاری، تاثیر متقابل خریدار و آنالیز. برای اتصال چنین اجزایی سه پل انتقال داده وجود دارد: stage data, deploy Results و customer Interaction. رابطه بین شرکت ها و پل های انتقال داده ها در شکل ۱ نشان داده شده است. سپس هر یک از اجزای معماری و پل هایی که این اجزا را به هم متصل می کنند در موردشان توضیح می دهیم.

در جزء تعریف داده تجاری، کاربر تجاری داده ها و فوق داده ها مرتبط با تجارتشان را تعریف می کند. این داده شامل اطلاعات تجاری (بطور مثال: تولیدات، لیست های قیمت و طبقه بندی) اطلاعات مفهومی (بطور مثال: الگوهای صفحه وب، تصویرها و چند رسانه ای و کالاها) و قوانین تجاری (بطور مثال: قوانین مفاهیم خصوصی شده، قوانین تبلیغ و قوانین چندفروشی و فروش زیاد) است. با استفاده از data mining دورنمایی به دست می آید که کلید جزء تعریف داده تجاری در واقع توانایی در تعریف یک مجموعه قوی از خصوصیات (فوق داده) برای هر نوع از داده می باشد.

بطور مثال تولیدات می توانند خصوصیتی همچون سایز، رنگ و در نظر گرفتن آن برای یک سن خاص و طبقه بندی شدن در یک مجموعه سلسله مراتبی همچون مردان و زنان و زیر مجموعه هایی که مثل کفش ها و لباس ها را داشته باشد. به عنوان یک نمونه دیگر

الگوهای یک صفحه وب می توانند ویژگی هایی داشته باشند که نشان دهد که آیا آنها تولیداتشان را نشان می دهند نتایج را جستجو می کنند، یا به عنوان قسمتی از فرایند بررسی (تسویه حساب کردن) مورد استفاده قرار می گیرند. داشتن مجموعه متغیری از ویژگی های موجود نه تنها برای data mining مفید است بلکه برای خصوصی سازی تجربه خریدار مفید است. عامل تاثیر متقابل خریدار بین خریداران و شرکت تجارت الکترونیکی رابطه برقرار میکند. اگر چه ما از نمونه یک وب سایت در طول این تحقیق استفاده کرده ایم، اصطلاح تاثیر متقابل خریدار بیشتر با هرگونه تاثیر بر خریداران در ارتباط است، این تاثیر متقابل می تواند در یک وب سایت رخ دهد (بطور مثال از طریق تلفن یا ایمیل)، کاربردهای بی سیمی یا حتی ساده ترین بخش شبکه فروش. برای تحلیل موثرتر همه این منابع داده ها، یک جمع کننده داده مورد نیاز است که جزء تلفیقی از اجزای تاثیر خریدار باشد. برای هماهنگی بیشتر، جمع کننده داده نه تنها باید مطالعات فروش داده را ثبت کند بلکه باید انواع دیگر تاثیرات متقابل خریدار از جمله تعداد بازدیدکنندگان صفحه وب را برای یک وب سایت ثبت کند. جزئیات بیشتر از معماری تلفیقی به منظور ایجاد یک وب سایت در بخش سوم توصیف شده است. برای اشاره کردن به سودمندی (خدمات رفاهی) این جمع کننده داده، بیایید یک نمونه از یک شرکت

تجاری که میزان تاثیر تبلیغات بانر شبکه خود را با دیگر سایتها انجام می دهد تا خریداران را به سمت سایت خودش جذب بکند، در نظر بگیریم. قیمت یک آگهی بانر شبکه معمولا بر اساس دفعاتی که بازدیدکنندگان وب سایت بر روی یک بانر تبلیغاتی و در طی دوره خاصی از زمان کلیک می کنند تعیین می شود یعنی هر بار که یک بازدیدکننده صفحه بر روی بانر تبلیغات کلیک می کند باید هزینه ای پرداخت شود. بسیاری از شرکتهای تجارت الکترونیکی میزان تاثیر تبلیغات بانر شبکه شان را با همین واحد اندازه گیری یعنی تعداد دفعات کلیک اندازه گیری می کنند و بنابراین در محاسبه میزان خرید هر فرد دچار مشکل می شوند. اگر هدف فروش بیشتر محصولات است پس سایت نیاز دارد که بیشتر خریداران را جذب کند تا اینکه دنبال جستجوگرها باشد. گزارش اخیر forrester نشان می دهد که « استفاده کردن از تعداد بازدیدهای صفحه به منظور قضاوت در مورد موفقیت یک سایت همانند ارزیابی یک اجرای موسیقی براساس بلندی صدای آن است.» در عمل ما می بینیم که نسبت خریدهای صورت گرفته به تعداد دفعات کلیک در حدود ۲۰ فاکتور در تبلیغات بانر شبکه ای یک شرکت متغیر است. یک تبلیغ پنج مرتبه تولید می شود همچون دیگر تبلیغات در زمینه فروش، بنابراین تعداد دفعات کلیک در تبلیغات پیشین



حدود  $\frac{1}{15}$  جریان کلیک بعدی نیست. توانایی اندازه گیری این نوع ارتباط نیازمند استفاده

از ترکیب منابع چندگانه داده می باشد.

جزء تحلیلی یک محیط تلفیقی برای حمایت تصمیم سودمندسازی انتقال داده ها، گزارش

کردن، الگوریتم های data mining تجسم وسایل OLAP است. در دسترس بودن

(غنی بودن) فوق داده های موجود داد های باعث شده است که جزء تحلیل گر دارای فواید

قابل توجهی در وسایل حامی تصمیمات افقی باشد، هم در زمینه قدرت و هم استفاده از

آن. بطور مثال به طور خودکار سیستم، نوع هر یک از ویژگی ها را می داند که شامل ارزش

ویژگی مجزا سفارش داده شده است یا مجموعه ای از ویژگی های دامنه دار روانه شده و

توضیحات متنی. برای یک وب سایت سیستم می داند که هر خریداری نوبت های شبکه

ای دارد و هر کدام از این نوبت های شبکه ای شامل تعداد بازدیدهای صفحه و سفارشات

خواهد بود. این مسأله را راحت تر می کند تا بطور اتوماتیک بتواند آمارهای مربوط به

خریداران، نوبت ها و تعداد بازدیدهای صفحه وب سفارش ها را محاسبه کند. ما تحلیل

جزء تلفیقی را در بخش ۴ بطور کامل شرح خواهیم داد.

پل data stage جزء تعریف داده تجاری را به جزء تاثیر متقابل خریدار وصل می کند.

این پل داده و فوق داده را در جزء تاثیر متقابل خریدار منتقل می کند (سفارش می دهد).

داشتن یک فرایند اجرایی چندین مزایا دارد، داشتن توانایی برای آزمایش تغییرات قبل از آنکه آنها را در محصولات اجرا کنیم، تغییرات بین فرمتهای داده ها را باعث می شود و بین دو جزء کپی کردن را برای سودمندی بیشتر سبب می شود، معاملات تجارت الکترونیکی را قادر می سازد تا فرمت های طلایی داشته باشند.

پل انبار داده ها جزء تاثیر متقابل خریدار را با جزء تحلیل گر بهم متصل می کند. این پل، اطلاعات جمع آوری شده در جزء تاثیر متقابل خریدار را به جزء تحلیل گر منتقل می کند و یک محل انبار اطلاعات جدید را برای اهداف تحلیلی ایجاد می کند. پل انبار داده ها همچنین همه داده های تجاری تعریف شده در جزء تعریف داده های تجاری (که به جزء تاثیر خریدار منتقل شده بود با استفاده از پل data stage). جمع کننده داده در جزء تاثیر خریدار معمولاً در یک سیستم پردازشگر معاملات on line یا OLTP که معمولاً طراحی شده است برای استفاده از روابط تکنیک های مدل سازی سیستم های OLTP تنظیم شده اند برای جا به جایی مناسب تعداد زیادی از update های کوچک و پرس و جو های کوتاه. برای اجرای یک معامله یک تجارت الکترونیکی این امر منتقدانه است اما برای تحلیل کردن [ ۴ و ۵ ] که معمولاً نیامند، اسکن های کاملی از چندین میز بسیار بزرگ و یک طرح ستاره مانند که کاربران تجاری بتوانند آنرا بفهمند، مناسب نیست. برای

data minig ما نیاز داریم که یک انبار ذخیره داده بسازیم و از تکنیک های مدل سازی چند بعدی استفاده کنیم. هم طرح انبار ذخیره داده ها و هم انتقال داده ها از سیستم OLTP به سیستم انبار ذخیره داده ها کاری بسیار پیچیده و وقت گیر می باشد. ساختن یک سازه از انبار ذخیره داده ها به عنوان یک بخش سازنده معماری بطور قابل توجهی پیچیدگی این وظایف (کارها) را کاهش می دهد. علاوه بر ETL معمولی (به عنوان، گزیده، انتقال و ذخیره) از لحاظ عملی پل واردات و تلفیق داده ها از سیستم های خروجی و تامین کنندگان تشکلی داده ها حمایت می کند(بطور مثال Acxiom). از آنجائیکه طرح سیستم OLTP توسط معمار کنترل می شود ما می توانیم بصورت خودکار طرح OLTP را به یک طرح ستاره ای چند بعدی تبدیل بکنیم که به منظور بهینه سازی تحلیل صورت گرفته است. آخرین پل deploy Result است که کلیدی را برای بستن حلقه و نتایج تحلیلی به کار می اندازد. این پل توانایی انتقال الگوها، امتیازها، نتایج و ویژگی های تازه ساخت را دارد. بازگشت انتقال داده ها را به Business Data Definition تعریف داده های تجاری و اجزای customer Interaction (عامل تاثیر متقابل خریدار را به کار می گیرد تا در قوانین تجاری برای خصوصی سازی استفاده کند بطور مثال خریداران بر اساس تمایل در قبول فروش چند گانه ) امتیاز داده می شوند و

سایت می تواند بر اساس این امتیازها شخصیت پیدا کند. این مرحله سخت ترین بخش از فرایند کشف آگاهی برای اجرا در یک سیستم غیر مجتمع می باشد. اگر چه فوق داده مشترک در کنار این سه جزء به این معنی است که نتایج مستقیماً می توانند در داده منعکس شوند در داده ای که کمپانی های تجارت الکترونیکی آنرا تعریف می کند.

### جمع کننده داده :

این بخش اجزاء جمع آوری داده معماری مورد نظر را توصیف می کند این جزء معاملات خریداران را ثبت می کند (بطور مثال خریدهها و پس دادن کالاها) و جریانهای (بطور مثال، جریانهای کلیک). از آنجائیکه جزء جمع کننده داده بخشی از تماس هر خریداری است (بطور مثال، وب سایت کاربردهای خدماتی خریدار و وسایل بی سیم) در این بخش ما به تفصیل جمع آوری داده ها در وب سایت خواهیم پرداخت. اغلب مفاهیم و تکنیک های ذکر شده در این بخش براحتی قابل تعمیم دادن در دیگر تماسهای خریداران می باشد.

### ثبت جریان کلیک :

اغلب معماری های تجارت الکترونیک براساس ثبت های وب سرور (خدمات وب) یا واریسی کننده بسته اطلاعاتی به عنوان یک منبع برای داده های جریان کلیک قابل تکیه کردن می باشد، از آنجائیکه این دو سیستم هر دو به دلیل غیر مزاحم بودن، مزیتی دارند که

سبب می شود تا از آنها برای قفل کردن هر کاربردی استفاده شود. در ثبت رویدادهای سطح بالا و کمبود توانایی در بهره برداری از فوق داده های موجود در این کاربرد عاجز است. یک ثبت شبکه معمولی شامل داده هایی نظیر صفحه در خواستی، زمان درخواست، آدرس HTTP (پروتکل انتقال ابر متن) خریدار و غیره برای هر درخواست خدمات شبکه می باشد. برای هر صفحه از طریق خدمات شبکه درخواست میشود، تعداد زیادی از درخواستها برای تصاویر و سایر مطالب روی صفحه. از آنجاییکه تمامی اینها در web server log ثبت شده اند، اغلب داده ها در log مربوط می شود به درخواستهای فایل های تصویری که اغلب برای تحلیل مفید هستند و معمولا Filter out شده اند. تمامی این درخواستها باید قبل از استفاده کردن از وبلاگ حذف شوند. بدلیل طبیعت فاقد تابعیت HTTP (پروتکل انتقال ابر متن) هر درخواست در یک وبلاگ از هر درخواست دیگری مستقل و جدا خواهد بود بنابراین شناسایی کاربران بسیار دشوار خواهد بود و یا اینکه نوبت کاربر از این داده [ ۶۷۸ و ۹ ] را بشناسیم. از آنجائیکه وبلاگ ها تنها شامل نام صفحه ای هستند که درخواست شده است بنابراین این نام های صفحات باید به مفاهیم، محصولات و غیره بر روی صفحه تبدیل گردند. این مشکل با معرفی مفاهیم دینامیک که در آنجا یک صفحه می تواند مورد استفاده قرار بگیرد برای نمایش میزان متفاوتی برای هر کاربر، بسیار

پیچیده می شود، در این حالت جزئیات محتوا که در یک صفحه شبکه نشان داده شده اند در یک وبلاگ قابل دسترسی نیست. مکانیزم مورد استفاده برای فرستادن داده مورد درخواستی به خدمات دهنده همچنین بر اطلاعات در وبلاگ تاثیر می گذارد. اگر مرورگر یک درخواست را بفرستد در حالیکه از روش post استفاده کند سپس پارامترهای ورودی برای این درخواست قابل ضبط کردن در وبلاگ نیستند.

واریسی کننده های بسته اطلاعاتی سعی می کنند که داده مشابهی را بوسیله نگاه کردن به داده در روی سیم جمع آوری کنند، اگر چه واریسی کننده بسته اطلاعاتی می تواند داده های بیشتری را نسبت به آنچه که در وبلاگ موجود است، نگاه کند اما آنها همچنان مشکلاتی دارند در زمینه شناسایی کاربران ( بطور مثال ورود به سیستم مشابه اما با دو دستگاه متفاوت) و نوبتها. همچنین راههای گوناگونی که وب سایتها بر اساس آنها طراحی شده اند سبب شده که راههای استخراج اطلاعات تجاری منطقی بوسیله نگاه کردن در جریان داده ها در طول یک سیم، بسیار دشوار گردد. یکی دیگر از مسائل پیچیده این است که واریسی کنندگان بسته اطلاعاتی می توانند داده را در محیط هایی از سایت که به منظور انتقال امن کد گذاری شده اند و همچنین از لحاظ کار کردن با سایتها مشکل دارند یا محیط هایی از یک سایت که از SSL استفاده می کنند را ببینند. چنین محیط هایی از

یک سایت بسیار مهم و اساسی هستند برای تحلیلی که شامل نظارت و فرم هایی که شامل داده های شخصی هستند. در بسیاری از سایتهای مالی همچون بانکها، کل سایت امن است بدین نحو که واریسی کنندگان بسته اطلاعاتی که داده های رمزی را کنترل می کنند، کور کرده و کاملاً غیر قابل استفاده می سازند بنابراین این واریسی کنندگان قبل از دسترسی به رمزها باید به اطلاعات دسترسی داشته باشند که این کار، تلفیق آنها را پیچیده تر می کند. گردآوری داده در لایه خدمات دهنده کاربردی به طور کاملاً موثری می تواند همه این مشکلات را حل کند. از آنجائیکه خدمات دهنده کاربردی اجزا را تامین می کند ( بطور مثال تصاویر، محصولات و مقالات). آگاهی شامل جزئیاتی را اجزا مورد نظر دارد. زمانی که اجزاء بصورت دینامیک تولید می شوند یا کدگذاری می شوند برای انتقال با استفاده از SSL، این امر قابل مشاهده است. خدمات دهندگان کاربردی از cookies (یا URL که در غیاب cookies کدگذاری می شود) برای نگه داری یک شیار از یک نوبت کاربر استفاده می شود بنابراین نوبت سازی جریان کلیک پیش افتاده است. از آنجائیکه خدمات دهنده کاربردی همچنین شیار کاربر را نگه می دارد از مکانیسم های ثبت یا cookies استفاده می کند، همکاری جریان کلیک را با یک بازدیدکننده ویژه آسان است. خدمات دهنده کاربردی همچنین می تواند طراحی شود برای نگه داشتن

Track اطلاعات غایب در لاگ های خدمات دهنده شبکه که شامل صفحاتی می شود که ناتمام گذاشته شده اند (کاربر دکمه توقف را فشار می دهد در حالیکه صفحه دانلود شده است)، وقت محلی کاربر، سرعت اتصالات کاربر و اگر کاربر cookies را خاموش کرده باشد این روش از گردآوری داده های جریان کلیک در هر دو زمینه وب لاگها و، واریسی کننده های اطلاعاتی توانسته پیشرفتهای قابل توجهی داشته باشد.

### ثبت رخدادهای تجاری : Business Event Logging

گردآوری داده های جریان کلیک از خدمات دهنده کاربردی بسیار قوی و جالب است : اگر جریان نگرش قابل توجه با نگاه کردن به در زیر مجموعه هاب درخواستها به عنوان یک رخداد منطقی یا اپیزود [ ۱۰ و ۶ ] می تواند بدست بیاید. اما این مجموعه از درخواستها را رخدادهای تجاری می نامیم. رخدادهای تجاری همچنین برای توصیف کارکردهای قابل توجه کاربر، مثل فرستادن یک ایمیل یا جستجو کردن [ ۲ ] مورد استفاده قرار گیرند. از آنجائیکه سرویس دهنده کاربردی باید مطالب نشست کاربران و داده های مرتبط را حفظ کند. بنابراین سرویس دهنده کاربردی یک انتخاب منطقی برای ثبت این رخدادهای تجاری می باشد. رخدادهای تجاری می توانند برای ترک گذاری چیزهایی مثل اجزای چرخ دستی های مخصوص خرید رها شده استفاده شوند بطوریکه بسیار مشکل خواهد



بود که برای ترک گذاری تنها از داده های جریان کلیک استفاده کنیم. رخدادهای تجاری همچنین کمک می کند به فروشندگان که به بالای صفحه نگاه کنند برای نسبت های تبدیل کوچک [ ۱۱ ] یک نسبت تبدیل کوچک برای هر مرحله از فرایند خرید به عنوان کسری از محصولاتی که بطور موفقیت آمیز به مرحله بعدی از فرایند خرید منتقل شده است، تعریف میشود. دو مثال از اینها شامل کسر نمایش محصولات که سبب میشود که محصول به چرخ دستی خرید اضافه شود و نسبت محصولات در چرخ دستی خرید که بطور موفقیت آمیز از هر فاز فرایند نظارت گذاشته است. از این رو دیدگاه یکپارچه موردنظر در این معماری به فروشندگان این توانایی را می دهد که به تماشای مستقیم منظره محصول، نماهای اجزا و فروش محصولات بپردازند یک توانایی که از تعداد دفعات کلیک بازدیدکنندگان و نماهای صفحه بسیار قوی تر است. بعضی از رخدادهای تجارت جالب که به تحلیل کمک می کنند در بالا ذکر شده و توسط معماری حمایت می شوند بدین قرار است :

• اضافه کردن / جابه جایی قسمتی به / یا از چرخ دستی خرید

• نظارت اولیه

• نظارت نهایی

- جستجوی رخداد

- ثبت رخداد

کلمات کلیدی جستجو و تعداد نتایج برای هر یک از این جستجوها که می تواند ثبت شود با جستجوی رخدادهایی که به فروشندگان نسبت به علاقه بازدیدکنندگانشان و تاثیر مکانیسم جستجو کردن نگرش موثرتری داشته باشند.

اندازه گیری میزان موفقیت خصوصی سازی :

معماری همچنین یک موتور Rules را پشتیبانی می کند که بر روی وب سایت برای خصوصی سازی بکار می افتد. Rules همچنین برای پیشنهاد تبلیغات به ملاقات کنندگان، نمایش محصولاتی خاص یا اجزایی خاص برای یک ملاقات کننده خاص و غیره استفاده می شود. یک رخداد تجاری هر زمان که یک قاعده در خصوصی سازی مورد استفاده قرار می گیرد، می تواند جمع آوری گردد. یا جفت شدن چرخ دستی خرید یا رخدادهای نظارتی، می تواند یک پیش بینی بسیار خوبی را از تاثیر هر قاعده بگوید. معماری همچنین می تواند از گروههای کنترل استفاده کند بنابراین در اینصورت قاعده خصوصی سازی تنها برای کسر خاصی از ملاقات کنندگان فعال می شود. این امر به تحلیل گران کمک می کند که مستقیماً به فروشها یا نتایج مربوط به بازدیدکنندگان نگاه کنند

زمانی که قواعد چه فعال و چه غیر فعال باشد. تکنیکهای گرد آوری داده های مشابه میتوانند برای همه تماسهای خریداران از جمله جایگزین سرویس خریدار، کاربرهای سیمی و غیره مورد استفاده قرار بگیرند. جمع آوری داده های صحیح برای یک تحلیل موثر از یک عملیات تجارت الکترونیکی منتقدانه است.

### تحلیل :

این بخش در ارتباط با اجزای تحلیلی معماری ما توضیح میدهد. ما با یک بحث از انتقال داده ها شروع کردیم که پس از آن در ارتباط با تکنیک های تحلیلی شامل گزارش، الگوریتم های data minig، تجسم و OLAP مواجه شدیم. محل انبار داده ها یک منبع داده از تحلیل ها در معماری ما می باشد. اگر چه الگوسازی چند بعدی معمولاً برای تحلیل یک پیش نیاز است اما تجربه ما نشان می دهد که بسیاری از تحلیلها نیازمند انتقال داده های اضافی می باشند که داده را به شکل هایی تابع پذیر تغییر دهند نسبت data minig همانطور که در قبل ذکر کردیم کاربر تجاری می تواند محصول، تبلیغ و سلسله مراتب طبقه بندی را تعریف کند در جزء تعریف داده تجاری شکل ۲ یک نمونه ساده از یک سلسله محصول است این اطلاعات سلسله مراتبی برای تحلیل بسیار قابل ارزش است اما تنها تعداد کمی از الگوریتمهای data minig موجود می توانند این را مستقیماً

استفاده کنند. بنابراین ما به انتقال داده ها نیاز داریم تا این اطلاعات را به شکلی تغییر بدسیم که توسط الگوریتم های data minig قابل استفاده باشند. یک راه حل ممکن این است که یک ستون را اضافه کنیم تا نشان دهد که شی در زیر کدام قسمت از سلسله مراتب قرار می گیرد بیایید از سلسله مراتب محصولات نشان داده شده در شکل به عنوان یک نمونه استفاده کنیم. برای هر خط دستور یا صفحه درخواست شامل یک محصول SKU (سهم، نگه داری، بخش).

این انتقال یک ستون bolian ایجاد می کند که با هر یک از بخش های انتخاب شده در سلسله مراتب در ارتباط است. این نشان می دهد که آیا این محصول که توسط یک کد جایگزین شده، تعلق دارد یا نه؟ شکل ۳ این عملیات را نشان می دهد.

از آنجائیکه خریداران مهمترین نگرانی هر تجارت الکترونیکی هستند اکثر تحلیل های data mining در سطح خریدار صورت می گیرد. به این دلیل هر یادداشت از داده در آخرین مرحله یک تحلیل جای می گیرد امضاء خریدار می باشد که در برگزیده همه اطلاعات در مورد خریدار می باشد. اگرچه اکثریت داده در انبار ذخیره داده ها در دیگر سطوح مثل سطوح سفارش دهنده، سطح خط سفارش و سطح صفحه درخواست قرار دارد. هر فروشنده ای ممکن است ردیفهای چندگانه ای در این سطوح داشته باشد. به منظور

مناسب ساختن این اطلاعات دقیق برای تحلیل در سطح خریدار، و مجموعه انتقالها ضروری به نظر می رسد. اینجا چند نمونه از ویژگی هایی که ما آنها را مناسب یافته ایم وجود دارد.

- چند درصد از هر یک از سفارشات خریداران از کارت اعتباری VISA استفاده کرده اند؟  
- چه مقدار پول هر خریدار برای کتابها خرج می کند؟

- چه مقدار هزینه کل همه خریداران از هزینه خانم ها بیشتر است؟

- مبلغ کلی پنج خرید اخیر هر یک از خریداران که بیش از ۳۰ دلار باشد، چه مقدار است؟

- هر خریدار هر چند وقت یکبار خرید می کند؟

- آخرین خرید هر یک از خریداران که شامل روزهای گذشته از تاریخ آن، چیست؟

این ویژگی ها برای ساختن بسیار سخت هستند و استفاده از نظریات استاندارد SQL نیازمند انتقالات اجماعی قدرتمندی هستند. ما به این نتیجه رسیدیم که این داده ها و ستونهای زمان، اطلاعات مفیدی را که الگوهای مهمی را آشکار می کنند، تبدیل می کنند.

اگرچه داده های معمولی و شکل زمانی شامل سال، ماه، روز، ساعت، دقیقه و ثانیه معمولاً

توسط الگوریتم های Data mining پشتیبانی نمی شوند. اغلب الگوهای مرتبط با داده

ها و زمانی مستقیماً از این شکل شناخته نمی شوند، به منظور کشف الگوهایی، داده ها و زمان را ساده تر به کار می برند، ما نیازمند انتقال های هستیم که بتوانند تفاوت بین زمان داده ها را حساب کنند (بطور مثال داده امری و داده ship) و ایجاد ویژگی هایی جدیدی که جایگزینی روز هفته، روز ماه، هفته، ماه، سه ماه از سال، سال و غیره. از داده و خصوصیات زمان باشند.

بر اساس مطالب ذکر شده فوق، معماری برای پشتیبانی یک سری تغییر شکل فراوان طراحی شده است. ما پی بردیم که تغییر شکل شامل: ایجاد خصوصیات جدید و افزایش خصوصیات سلسله وار، اجماع، فیلتر، نمونه، ستونهای حذفی و امتیاز برای آسانتر کردن تحلیل مناسب است با توجه به توضیح تغییر شکلها بیائید در ارتباط با ابزار تحلیلی بحث کنید. گزارش اصلی یک نیاز مهم برای تجارت الکترونیکی است. بر اساس گزارشات به دست آمده، کاربران تجاری می توانند پی ببرند که چگونه یک وب سایت در سطوح مختلف و بر اساس دیدگاههای مختلف، کار می کند. سوالهای مثال که با استفاده از گزارشات جواب داده می شوند بدین قرار است:

- محصولات پر فروش کدام هستند؟

- محصولات کم فروش کدام هستند؟

- www.kandooon.com
- صفحاتی که بیشترین بیننده را داشته اند، کدام اند؟
  - جستجوهای که بیشترین شکست را داشته اند؟
  - میزان گفتگوها به وسیله Brand چقدر است؟
  - میزان گستردگی مرورگرهای شبکه چگونه است؟
  - بیشترین قسمتی که مورد ویزیت قرار گرفته بر اساس مبلغ هر ویزیت کدام است؟
  - محصولاتی که بیشترین آمار رکود را دارند، کدام هستند؟
- تجربه به ما نشان داده که برخی از سوالهای گزارش، مانند دو سوال آخر بدون معماری تلفیقی که هم جریانهای رخداد و هم داده های فروش را ضبط کند، بسیار سخت است. نسل مدل که از الگوریتمهای data paining استفاده می کنند، یک مفهوم کلیدی از معماری می باشد. این نسل، الگوهایی را در مورد خریداران، خریده های آنها، صفحات نمایش و غیره ارائه می دهد. به وسیله تولید مدل ها، و می توانیم به سوالات زیر پاسخ دهیم.
- افراد زیاد هزینه کننده چه خصوصیتی دارند؟
  - چه خصوصیتی در افراد خریدار باعث می شود که آنها کالای X را به کالای Y ترجیح دهند؟
- www.kandooon.com

- چه خصوصیتی، خریدارانی که انواع فروش CROSS و UP را بر می گزینند، مشخص می کند؟

- چه خصوصیتی خریدارانی را که سریع خرید می کنند، مشخص می کند؟

- چه دسته ملاقات کنندگان خاصی هستند که خرید نمی کنند؟

بر اساس تجربه ما، علاوه بر الگوریتم های اتوماتیکی **data mining**، وسایل تغییر مدل موثر برای پشتیبانی نگرش تجاری باید به کار روند. مدل ها بر حسب اینکه اتوماتیک تولید شوند و یا توسط تغییرات موثر حاصل شوند، بعدها می توان آنها را با تست داده ها، امتحان و یا ارزیابی کرد. هدف این است که کاربران تجاری، قبل از اینکه مدل هایشان را اجرا کنند و به حالت ستونی در آورند، آن را بفهمند. به طور مثال، پی بردیم که برای مدل های قاعده ای، اندازه گیریهای همچون اعتماد به نفس، روحیه و پشتیبانی در سطح قاعده فردی و سطح پیوسته فردی به علاوه برای صحت کلی مدل سودمند است. در تجربه ما، کاربردهای زیر برای اصلاح موثر یک مدل قاعده مناسب است:

- بتواند اجزاء را ببیند (به طور مثال، قسمت خریدار) به وسیله یک زیر مجموعه از قواعد یا یک زیر مجموعه از پیوست های یک قاعده.



- بتواند به طور غیر خودکار، یک رول قاعده را تغییر دهد به وسیله حذف کردن، اضافه

کردن یا تغییر یک قاعده یا پیوست فردی.

- به طور مثال، یک مدل قاعده هزینه کنندگان زیاد را پیش بینی می کند. با توجه به

مقادیر زیر:

اگر دلار ۸۰/۰۰۰ درآمد

۳۱ ← سن

میانگین دوره نوبت بین ۱۰ تا ۲۰/۱ دقیقه است و تاریخ این هزینه قبل از ۲۰۰۱/۱/۴

است.

سپس یک هزینه کننده زیاد.

ممکن است شما تعجب کنید که چرا تغییر سن به جای ۳۰ در ۳۱ رخ داده است و جدا

سازی میانگین مدت نوبتها در دقیقه ۲۰/۱ رخ می دهد به جای دقیقه ۲۰. چرا داده ایجاد

حساب در این قاعده اصلا به حساب نیامد؟ یک کاربر تجاری ممکن است بخواهد قاعده را

به شکل زیر تغییر دهد:

اگر دلار  $80/000 >$  درآمد

$30 \leq$  سن

میانگین دوره نوبت بین ۱۰ تا ۲۰ دقیقه است.

سپس هزینه کننده زیاد

اگرچه قبل از انجام کار، مهم است که ببینیم چگونه این مرحله، اندازه گیریهایی مثل (اعتماد به نفس، روحیه و حمایت) از این قاعده و مدل کلی قاعده را تغییر می دهند.

با علم به این مسئله که انسانها در تشخیص الگوها از داده های تجسمی بسیار خوب هستند. تجسم و ابزار OLAP می توانند به کاربران تجاری، بسیار کمک کنند تا بوسیله

ابزار گزارش مکمل و الگوریتمهای data mining، بتوانند نسبت به مشکلات تجاری آگاهی پیدا کنند. تجربه ما می گوید که ابزار تجسمی برای درک مدل های تولید شده،

عملیات وب سایت و خود داده ها بسیار کمک خواهد کرد. شکل ۴ یک نمونه از ابزار تجسمی را نمایش می دهد که به طور واضح نشان می دهد که خانم های بنی ۳۰ تا ۳۹

سال افراد بسیار هزینه کننده ای هستند (مربع بزرگ). همین حالت برای مردان بین ۴۰ تا ۴۹ سال است.

#### ۵- چالش ها (Challenges)

در این بخش ما چندین شکل چالش آور را بر اساس تجربه مان در استخراج داده های تجارت الکترونیکی است. سطح پیچیدگی این شکل ها با هم فرق می کند، اما به هر حال

هر کدام جایگزین یک محیط زندگی باشند، جایی که ما اعتقاد داریم که می توان در آنجا پیشرفت کرد. به جز دو چالش اولی، با چالش الگوریتمی data mining حل خواهد شد.

فراگیر ساختن مدل های data mining برای کاربران تجاری: Make Data Mining

### Models Comprehensible to Business Users

کاربران تجاری از تاجران گرفته که در طبقه بندی محصولات تصمیم می گیرند تا طراحان وب سایت ها را طراحی می کنند و یا بازاریابانی که تصمیم می گیرند که کجا برای تبلیغ دلار خرج کنند، همه نیاز دارند که نتایج data mining را بفهمند. گزارشات چکیده، ساده ترین راه برای درک و معمولاً قابل فراهم ساختن به ویژه حوزه های خاص افقی هستند. تجسم های ساده مثل نمودارهای ستونی، کاغذ شطرنجی دو بعدی که برای فهمیدن آسان است و می تواند اطلاعات بیشتری را فراهم سازد و الگوها را برجسته کند، به خصوص اگر با رنگ به کار رود. تعداد کمی از مدل های data mining برای فهمیدن آسان است. قواعد طبقه بندی ساده ترین حالتی است که با درختهای طبقه بندی دنبال

می شود. یک تجسم برای Nave-Bayes [۱۲]

برای کاربران تجاری به منظور فهمیدن تجربیات گذشته، نویسنده دوم آسان است.

چالش تعریف بیشتر انواع مدل (فضاهای فرضیه) و راههای ارائه آنها به کاربران تجاری چه مدل های رگرسیون را پیدا خواهیم کرد و چگونه می توانیم آنها را ارائه دهیم. (حتی فهمیدن رگرسیون بین خطی برای کاربران تجاری سخت است). چگونه می توانیم به طور مثال مدل های نزدیک ترین همسایه را ارائه دهیم؟ چگونه می توانیم نتایج الگوریتمهای قاعده همکاری را بدون تحت فشار قرار دادن کاربران دهها هزار قاعده ارائه دهیم. (یک نمونه خوب از این مشکل در کتاب [۱۳] Berry and linoff در ابتدای صفحه ۴۲۶ کتاب آورده شده است.

قابل دسترس ساختن تغییر شکل داده ها و ساختمان مدل برای کاربران تجاری:

توانایی برای پاسخ دادن به یک سوال که توسط یک کاربر تجاری ارائه می شود معمولاً نیازمند چند تغییر شکل داده و فهم تکنیکی ابزار می باشد. تجربه به ما نشان می دهد که حتی طراحی گزارشات تجاری و ابزارهای OLAP برای اکثر کاربران تجاری بسیار سخت است. دو راه حل معمولی شامل فراهم کردن قالبها (بطور مثال قالبهای گزارش، مکعب های OLAP و تغییر شکلهای ذکر شده برای فرایند می باشد). برای سوالهای معمول، چیزهایی که در فروشگاههای افقی خوب تعریف شده به خوبی کار می کنند و فراهم ساختن مهارت از طریق مشاوره یا یک سازمان خدمات رسانی چالش برای پیدا کردن

راههای است در اجازه دادن به کاربران تجاری که در نتیجه بتوانند خودشان خدمت دهی کنند.

پشتیبانی سطوح گرانولیتة چندگانه:

داده های گردآوری شده در یک وب سایت معمولی شامل یادداشتهایی در سطوح مختلف گرانولیتری می باشد:

صفحه های نمایش پایین ترین سطح هستند با ویژگی هایی همچون نمایش محصولات و طول مدت.

نوبتهایی شامل ویژگی هایی همچون استفاده مرورگر، زمان آغاز، سایت رجوع شونده و اطلاعات کوکی.

هر نوبت شامل صفحه نمایش های چندگانه می باشد.

خصوصیات خریدار شامل نام، آدرس و ویژگی های جمعیتی. هر خریدار ممکن است در نوبتهای چندگانه محسوب شود. استخراج در سطح صفحه نمایش به وسیله به هم پیوند

دادن همه نوبتها و خصوصیات خریدار، فرضیه اصلی در اکثر الگوریتم های *data mining*

به خصوص آن یادداشت هایی که بطور مستقل و همانندی پخش می شود را بر هم می

زند. اگر سعی داریم که یک مدل را بسازیم تا پیش بینی کنیم که چه کسانی صفحه X را

دیدن کرده اند و جو اتفاقا این صفحه را اغلب اوقات دیدن کرده است پس ممکن است ما به قاعده ای برسیم که اگر اولین اسم یک بازدیدکننده، جو باشد پس آنها حتما صفحه X را ملاقات کرده اند. قاعده یادداشتهای چندگانه ای (ملاقاتها) برای پشتیبانی آن، اما این امر به طور واضح نمی تواند فراتر از همان جو مورد نظر کلیت داده شود.

این مشکل بوسیله استخراج مشکلات دیگر در حوزه ارتباط از راه دور در ارتباط است [۱۴]. چالش اینست که الگوریتم هایی را طراحی کنیم که بتواند سطوح گرانبه چندانگانه را به درستی پشتیبانی کنند.

به کار بردن سلسله مراتبها:

محصولات معمولا در سلسله مراتبها طبقه بندی می شوند: skuها از محصولات گرفته شده اند آنهایی که از خانواده محصولات گرفته شده، آنهایی که از دسته ها گرفته شده اند و غیره. یک سلسله محصولات معمولا از سه تا هشت سطح عمق دارد. یک خریدار موارد sku را خریداری می کند اما کلی سازی در سطوح بالاتر به دست می آید (بطور مثال خانواده ها و طبقه بندیها) بعضی الگوریتم ها طراحی شده اند برای ویژگیهای درخت مانند [۱۵]، اما آنها در سلسله مراتبهای بزرگ محصول بالا نمی روند. پشتیبانی چنین

سلسله مراتبی بوسیله الگوریتم های data mining یک چالش است.

اندازه گیری بهتر، جابه جایی مقدار عظیمی از داده ها:

یاهو، ۴۶۵ میلیون بازدید کننده در روز دسامبر ۱۹۹۹ داشته است. چالش اینست که

تکنیک های مناسبی را پیدا کنیم (جدای از نمونه برداری). که این داده ها را به مقیاس

درآورد. آیا مجموعه هایی وجود دارد که بتوانند هنگامیکه داده جمع آوری می گردد اجرا

شوند.

پشتیبانی در الگوردهای رخدادهای بیرونی:

رخدادهای بیرونی همچون عملیاتهای بازاریابی (بطور مثال، تبلیغات و آگهی های رسانه

ای و سایت دوباره طراحی می کند الگوهای متفاوتی را برای داده. چالش اینست که بتوانیم

چنین رخدادهایی را الگودهی کنیم که بتوان الگوهای جدیدی را ایجا کرد که در طول

زمان جرقه هایی زده شود.

پشتیبانی تغییرات آهسته ابعاد:

جمعیت شناسی بازدیدکنندگان تغییر می کند: افراد ازدواج می کنند، بچه هایشان رشد

می کند، حقوقشان تغییر می کند. با این تغییرات، نیازهایشان که قبلا الگوگذاری شده بود

تغییر می کند. ویژگی های محصولات تغییر می کند:

انتخابهای جدید (بطور مثال رنگها) ممکن است بوجود بیاید، نوع بسته بندی کالاها یا طرح آنها تغییر کند و یا حتی کیفیت یک کالا ممکن است بهتر شود و یا بدتر شود. این خصوصیات که در طول زمان تغییر می کنند، معمولا اغلب به عنوان تغییرات آهسته ابعاد یاد می شوند [۴]. چالش این است که این تغییرات را دنبال کنیم و برای چنین تغییراتی در بحث تحلیلی، بخش پشتیبانی داشته باشیم.

شناسایی و robots و crawler ها:

Crow lers و روبات ها به طور قابل توجهی می توانند الگوهای جریان کلیک در یک وب سایت را تغییر دهند. به طور مثال اسم key note ([www.keynote.com](http://www.keynote.com))، ابزارهای اندازه گیری اجرای سایت را فراهم می سازد. روبات های key notes می تواند یک درخواست چند زمانه را در یک دقیقه تولید کند، ۲۴ ساعت روز، ۷ روز هفته، آمار مربوط به تعداد نوبتها، صفحه ها و صفحه های خروجی را (آخرین صفحه در هر دوره) را تحریف می کند.

موتورهای جستجو هدایت می کنند اسکن های عریض اولیه از سایت، تولید تعدادی درخواست در دوره کوتاه. جستجوگر اینترنت ۵ تنظیم اتوماتیک صفحه های شبکه را هنگامی که یک کاربر در آن ثبت می کند و زمانی که کامپیوتر بیکار است یا در یک برنامه



مشخص است را حمایت می کند. این جستجوگر همچنین، جستجوگری of line را که صفحات را از یک عمق مشخص از یک صفحه سود می کند را پشتیبانی می کند. این گزینش ها جریان کلیک و الگوهای جانبی را تولید می کند. شناسایی چنین روبات هایی برای فیلتر کردن جریان کلیک های آنها یک کار غیر بدیهی است بخصوص برای رباتهایی که تظاهر می کنند که کاربران واقعی هستند.

خلاصه:

ما یک معماری را مطرح کردیم که به طور موفقیت آمیزی data mining را سیستم تجارت الکترونیکی مجموعه سازی می کند. معماری ذکر شده شامل سه جزء مهم می باشند: تعریف داده تجاری، کنش خریدار و تحلیل که با استفاده از پل های انتقال داده متصل شده است. این مجموعه به طور موثری چند مشکل مهمی که با ابزار افقی data mining همراه است را حل می کند که شامل تلاش زیاد برای پردازش داده قبل از اینکه برای استخراج به کار رود و نتایج استخراج را عملی سازد. مجموعه محکم بین سه جزء معماری برای ساختار اتوماتیک یک انباره ذخیره در کنار سه جزئی که این ساختار را راحتتر می سازد و با یک سری الگوریتم های استخراج و ابزار تحلیلی (مثل تجسم، گزارش و OLAP) همچنین مناسب بودن فرایند کشف آگاهی را افزایش می دهد مجموعه محکم

و فوق داده مشترک همچنین باعث می شود که به صف درآوردن نتایج آسان شود و به طور موثری چرخه را پایان دهد. در پایان ما چندین چالش مشکل را ارائه کرده ایم که برای پیشرفت این معماری باید روی آنها بیشتر کار شود.

سپاسگذاری:

ما به امتنان از سایر اعضا تیم های data mining و تجسم شرکت نرم افزارهای در بلومارتینی و نویسنده اسناد سیندی هال تشکر می کنیم ما همچنین از سرویس گیرندگان تشکر می کنیم که داده هایشان را با ما در میان گذاشتند و به ما کمک کردند و معماری ما را ارتقاء بخشیدند و محصولات بلومارتینی را بهبود دادند.