

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



University of Kashan

دانشکده مهندسی برق و کامپیوتر

کاربرد بینایی کامپیوتر در تشخیص خودکار بیماری‌ها

Computer Vision Applications in Automated Disease Diagnosis

شماره دانشجویی: ۴۰۰۲۱۱۶۰۲۱۴

گردآورنده: سینا مظفری راد

گزارش نهایی درس روش پژوهش

مهندس یوسفان

خرداد ۱۴۰۳

چکیده

این نوشتار، ساختار پایگاه داده Dunnhumby را که از خرده‌فروشی انگلستان گردآوری شده بررسی می‌کند. ابتدا به تفصیل جداول کلیدی مانند transactions.csv, products.csv, households.csv و Marketcasualdata.csv. (Basket Analysis)، خوشه‌بندی مشتری (Customer Segmentation)، مدل‌سازی پیش‌بینی (Predictive Modeling) و ارزیابی کمپین‌های بازاریابی (Campaign Effectiveness) مطرح می‌گردد. ضمن اشاره به مزایا و چالش‌های پژوهشی، چشم‌انداز استفاده از دیتاست Dunnhumby در تحلیل رفتار مصرف‌کننده و توسعه راهکارهای داده‌محور در خرده‌فروشی ارائه خواهد شد.

واژگان کلیدی: Dunnhumby، خرده‌فروشی، تحلیل سبد خرید، خوشه‌بندی مشتریان، مدل‌سازی پیش‌بینی، داده‌کاوی

فهرست مطالب

۱ مقدمه

دیتاست Dunnhumby یکی از مراجع اساسی برای پژوهشگران حوزه خردهفروشی است. این پایگاه داده، داده‌های مرتبط با تراکنش‌های مشتریان، ویژگی‌های دموگرافیک خانوارها و اطلاعات محصولات را به صورت جامع ارائه می‌کند. تحلیل چنین اطلاعاتی می‌تواند الگوهای پنهان در رفتار مصرف‌کنندگان را آشکار سازد و در مدیریت ارتباط با مشتری، بهینه‌سازی سبد خرید و بهبود راهبردهای بازاریابی سودمند باشد.

۲ ساختار پایگاه داده و جداول کلیدی

در ادامه، فایل‌های اصلی این دیتاست همراه با ستون‌های مهم هر یک معرفی می‌شود.

۱.۲ جدول transactions.csv

تعریف و عملکرد: شامل سوابق خرید مشتریان، از جمله زمان و مقدار خرید هر محصول. ستون‌های اصلی:

$BASKET_ID$: $household_key$: □
•

□ DAY : شماره روز یا تاریخ خرید

□ $QUANTITYPRODUCT_ID$: تعداد کالا در هر تراکنش

$STORE_ID$: $SALES_VALUE$: □
•

$COUPON_DISC$: $RETAIL_DISC$: □
•

$TRANS_TIME$: $COUPON_MATCH_DISC$: □
•

□ $YEARWEEK_NO$: سال

۲.۲ جدول products.csv

تعریف و عملکرد: توصیف ویژگی‌های محصولات، شامل نام تجاری و گروه کالایی. ستون‌های اصلی:

MANUFACTURERPRODUCT_{ID} : سازنده یا تولیدکننده

DEPARTMENT : دپارتمان کالا (مثلاً نوشیدنی‌ها)

BRAND : برند کالا

SUB_{COMMODITY}_DESC : COMMODITY_DESC :

.

CURR_{SIZE}_{OFP}PRODUCT :

hh_{demographic.csv}

۳.۲ جدول households.csv یا

تعریف و عملکرد: ارائه اطلاعات دموگرافیک خانوار که برای مطالعه الگوهای خرید بر پایه سن، درآمد و ساختار خانوار کاربرد دارد. ستون‌های اصلی:

age_{band} : AGE_DESC_{household_{key}} :

.

income_{band} : INCOME_DESC_{marital_{status}} : MARITAL_STATUS_CODE :

.

HH_{COMP}_DESC : () home_{ownership} : HOMEOWNER_DESC :

.

KID_CATEGORY_DESC : household_{size} : HOUSEHOLD_SIZE_DESC :

STORE_ID :
PRODUCT_ID : ☐

•

✉ mailer: درج یا عدم درج در نامه تبلیغاتی

این جداول برای درک بهتر فرآیندهای بازاریابی و ارائهٔ کوپن در Dunnhumby به کار می‌روند. برای مثال، `coupon.csv` محصولاتی را مشخص می‌کند که تحت تخفیف قرار می‌گیرند و `coupon_redemptions.csv`

3.1 □□□□□ □□□ □□□□)Market Basket Analysis)

3.2 □□□□□□□ □□□□□□□)Predictive Modeling)

3

3.3 □□□□□□□□ □□□□□□□□)Customer Segmentation)

□□□□□□ □□□□□□□□□□ □□□□□ K-means □□□□ □□□□□□□□ □□□□□□□□
□□ □□□□ □□□□□□□□□□ □□□□ □□□□□□□□ □□□□□ □□□□□□ □□□□□□
□□□□□□ .□□□□□□□□

3.4 □□□□□ □□□ □□□□□)Time Series Analysis)

□□□□□□□ □□□□□ DAY□ WEEK_N□ YEAR □□ □□□□ transactions□ □□□□□
□□□□□□ □□□□□□□ □□□□ □□ □□□□□□□ □□□□ □□ □□□□□□ .□□□□□□□

3.5 □□□□□□□ □□□□□□□)Campaign Effectiveness)

□□□□□□ □□□□□ □□□□□ □□ □□□□□□□ □□ □□□□□□ □□ transactions □□□□
□□□□ □□□□□ □□□□□□ □□ □□□□□ □□ □□□□□ □□ □□ □□□□□ □□□□
□□□□□ □□□□□□□ .□□□□□

4 □□□□□□ □ □□□□□□□□ □□□□□□□

:□□□□□□

- □□□□□ □□□□□□□ □□ □□□□ □□□□□□□ □□ □□□□□□□ □□□□□□
- □□□□□□□□□□ □□□□□ □□ □□□□□□□ □□□□□ □□□□□□□□□□
- □□□□□□ □□□□□ □ □□□□□□ □□□□□ □□□□□□□□□□ Big Data
- □□□□□□□ □□□□□□□□□ □□ □□□□□□ □□□□ □□□□□□□□ □□□□□
□□□□ □□□□ □□□□□□

:□□□□□□□

- □□□□□ □□ □□□□□□□□ □□□□□ □□□ □□□□□□□ □□□ □□□□□□ □□□□
- □□□□□□□□□ □□□□□ □ □□□□□ □□□□□□□□ □ □□□□□ □□□□□□□

- □□□□□ □□□□ □□□□□ □□□□□□□ □ □□□□□□ □□□□□□□
- □□□□□□□ □□□□□□□ □ □□□□ □□ □□□□ □□□□ □□ □□□□□□□ □□□□□

5 □□□□□□□

□□□□□ **Dunnhumby** □□ □[□]□□□□ □□□□□□□ □□□□□□□ □□□□□ □□□□□
□ □□□□□□□ □□□□□□□□ □□□□□ □□□□□□□□□□ □□□□ □□□□□□□□
□□□□□□□□ □□□□□.□□□□□□ □□□□□□□□□□ □□□□ □□□□□ □□□ □□□□□
□□□□□□□□ □□□□□□□ □ □□□□□□ □□□□□□□□ □□□□ □□ □□□□□ □□□□□□□
□□ □□□ □□□□□□ □□□□ □□□□.□□□□□□ □□ □□□□ □□□□□□□□ □□□
□□□ □□□□ □□□□ □ □□□□ □□□□□ □□□□ □□□□□□ □[□]□□□□□□ □□□□□
□□ **Dunnhumby** □□□□□□□ □□□□ □□ □□□□□ □□□□□□□□ □□□□□□□□□□
□ □□□□□□ □□□□□□□□□□ □□□□□□□□□□□□.□□□□