

به نام خدا

## مقدمه

هدف سند (طراحی سطح پایین) ارائه منطق حل مساله به صورت گام به گام است.

طراحی کد برنامه واقعی برای داشبورد تجزیه و تحلیل فروش grocery آمازون است.

## ۱. معرفی پروژه

این یک پروژه در مورد مدیریت فروش مواد غذایی است. به دنبال داده کاوی، تحلیل داده و ارتباط میان خریداران و میزان محصولات مورد فروش است.

## ۲. بیان نیازمندی

مدیریت فروش برای پاسخگویی به رقابت فزاینده و نیاز به بهبود روش های توزیع برای کاهش هزینه و افزایش سود اهمیت پیدا کرده است. امروزه مدیریت فروش مهمترین وظیفه در یک شرکت تجاری است.

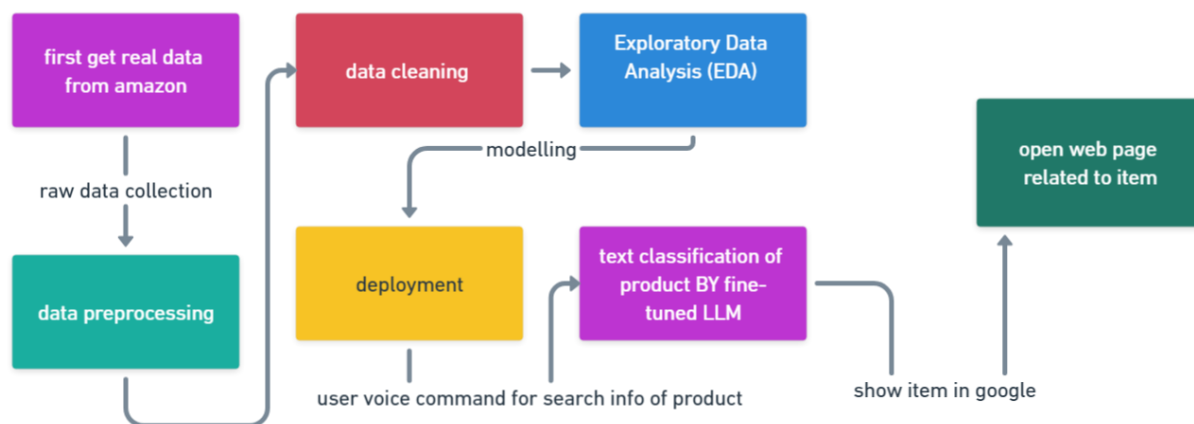
برخی از مجموعه داده های آمازون را استخراج شده است -

روند فروش به صورت ماهانه، سالانه، پیدا شده است و معیارها و عوامل کلیدی را بیابید و روابط معنادار بین ویژگی ها را نشان داده شده است.

## ۳. اطاعات دیتاست

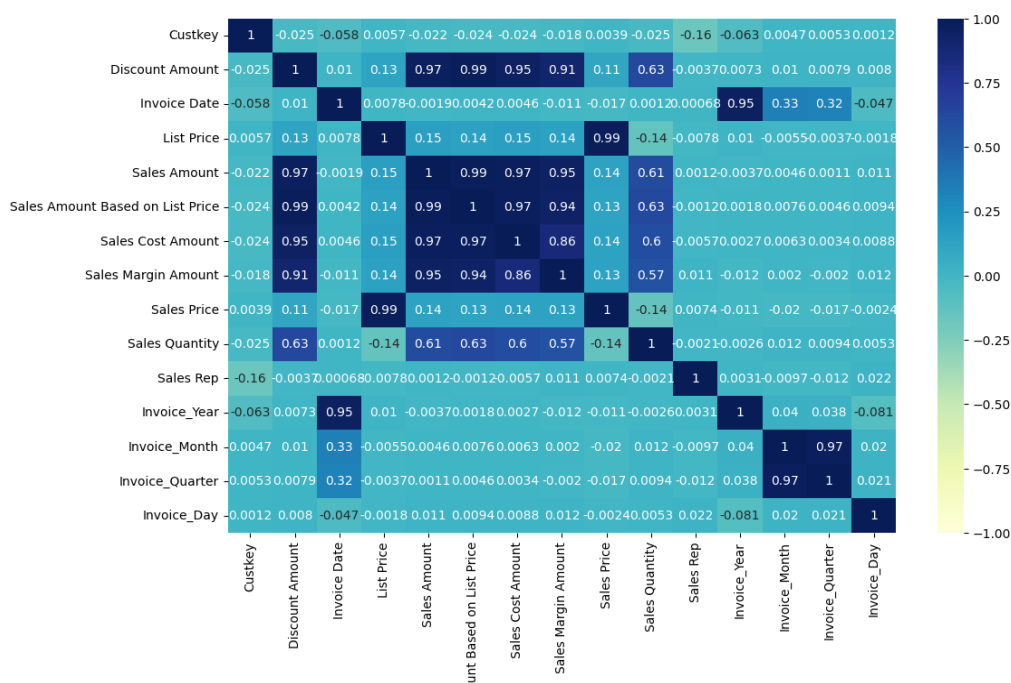
تاریخ فاکتور: روزی که فاکتور تولید شد.	Invoice Date
مجموع تخفیف ارائه شده برای هر کالا.	Discount Amount
مقدار فروش: مجموع قیمت فروش یک کالا.	Sales Amount
میزان حاشیه فروش (سود): میزان حاشیه فروش تفاوت هزینه فروش است	Sales Margin Amount (Profits)
مبلغ و مبلغ فروش. مقدار هزینه فروش: کل قیمت تمام شده هر کالا.	Sales Cost Amount:
قیمت فروش: قیمت فروش هر کالای خاص.	Sales Price
لیست قیمت: قیمت پایه یک کالا همانطور که در لیست قیمت منتشر شده است.	List Price
نماینده فروش: شخصی که شغلش فروش محصولات یا خدمات برای یک شرکت است.	Sales Rep
واحد اندازه گیری	U/M
یک شماره منحصر به فرد در فاکتور است که برای ارجاع به مشتریان استفاده می شود. شماره حساب کاربری	CustKey

## ۴. معماری

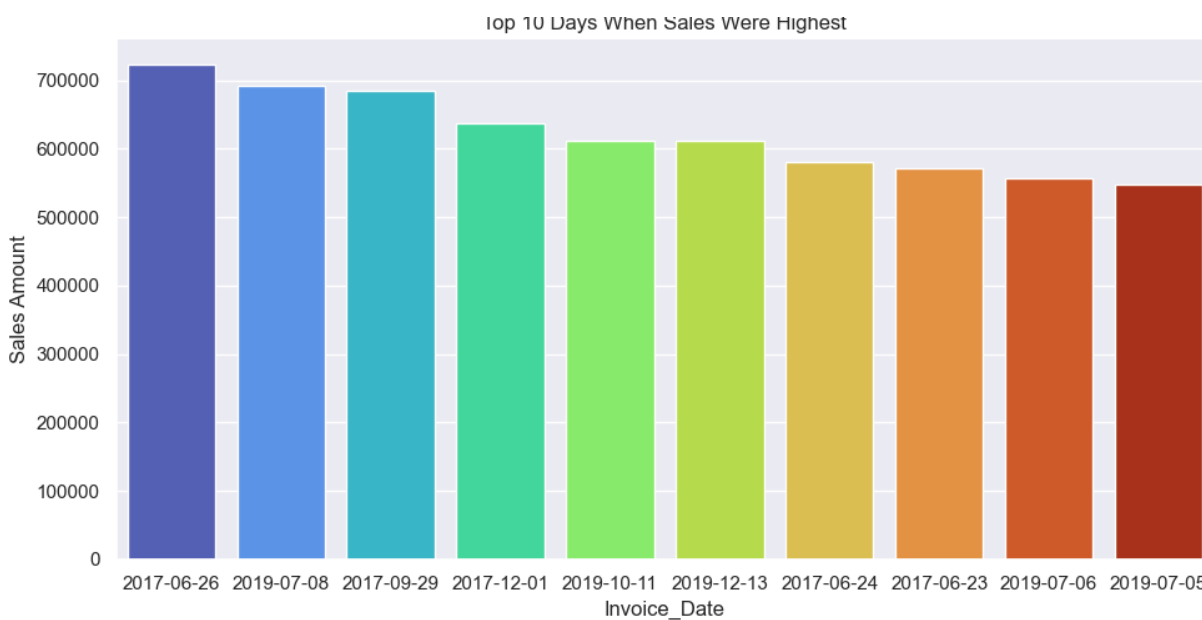


در ابتدا یک دیتا ست از آمازون حاوی ۶۵۲۸۱ داده استخراج می کنیم.

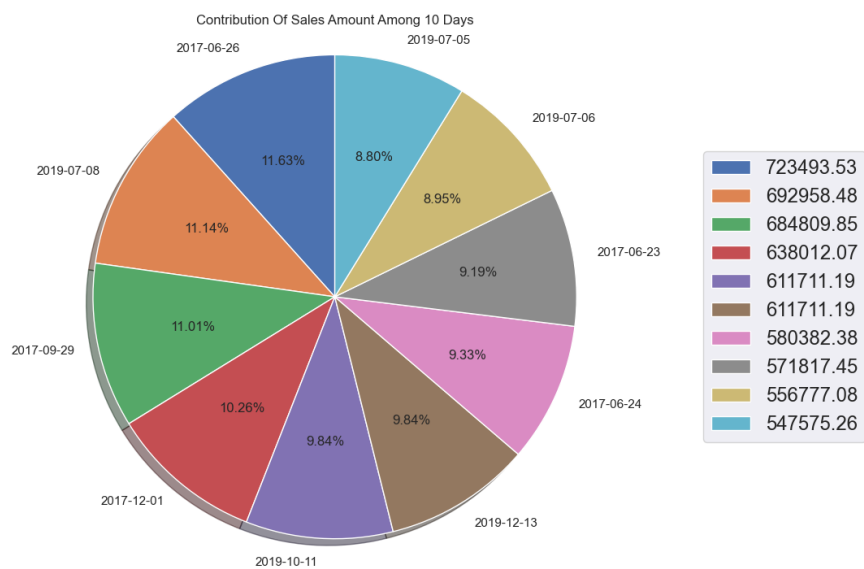
سپس با استفاده از تحلیل همبستگی پیرسون و همبستگی میان ویژگی ها و فیچرهای دیتا ست متوجه می شویم. هر چه میزان تخفیفات بیشتر باشد میزان فروش هم بالاتر می رود. چنانکه ۱۰ ایتم پر فروش دارای بیشترین تخفیف بوده اند. حتی با ارائه تخفیفات کم میتوان به سود قابل توجهی دست یافت. در تصاویر به ترتیب میزان همبستگی فیچرها و مشخص شدن میزان فروش به میزان تخفیف اجناس است.



در تصویر دوم میزان تاریخ ها با فروش بالای اجناس را نشان می دهد که بیانگر آنست که در طول دوره های زمانی خاصی فارغ از وجود تخفیف فروش اجناس بالا می رود. یک نتیجه محتمل آن است که اعمال تخفیفات بر اجنای در این دوره های زمانی میتواند فروش و سود حاصله را چند برابر کند.



تصویر سوم دوره ای ۱۰ روزه پر فروش را نشان میدهد.



سپس لیستی از ۱۰ محصول پر تخفیف به کاربر نشان داده می شود.

-----  
Top 10 unique most discounted items:

	Item Number	Item	Discount Amount	Sales Amount
12178	28500	Ebony Squash	191196.5532	200728.17
14451	380762	Fabulous Cranberry Juice	141552.9000	110213.10
15480	37570	Walrus White Zinfandel Wine	135794.2000	114268.80
15226	26502	Bravo Large Canned Shrimp	119797.0000	124491.20
17058	60776	High Top Sweet Onion	100197.8900	85678.71
64794	62055	Big Time Frozen Peas	98168.8640	82607.62
10052	64941	Ebony Shitake Mushrooms	89516.0400	75326.40
630	265072	Fabulous Mango Drink	88843.4300	75969.57
21752	116908	Golden Frozen Corn	87513.6234	78477.35
6857	265062	Fabulous Berry Juice	78786.8100	67370.19

-----

چندمین محصول پر تخفیف را می خواهید؟ لطفاً بین 1 تا 10 انتخاب کنید و سپس enter را بزنید

با انتخاب کاربر به طور مثال شماره ۱ جنس شماره ۱ برای کاربر ترجمه می شود.

-----  
چندمین محصول پر تخفیف را می خواهید؟ لطفاً بین 1 تا 10 انتخاب کنید و سپس enter را بزنید

-----  
Original: Squash

Translated: کدو

سپس کاربر می گوید کدو چیست؟

صدای کاربر اگر ضبط کند و در مسیر برنامه قرار دهد و اگر فایلی وجود نداشته باشد میکروفن باز شده با استفاده از api google صدای کاربر تشخیص داده می شود و به متن تبدیل می گردد.

میکروفن فعال است. لطفا صحبت کنید:

فکر می کنم گفتید: - کدو چیست

Process finished with exit code 0

حال متن به **parsbert** جهت کلاسیفیکیشن داده می شود. اما چون کدو در دیتا **parsbert** وجود ندارد. نیاز داریم مدل را **fine tune** کنیم.

برای اینکار ابتدا دو دیتا ست تهیه میکنیم برای آموزش و آزمایش.

content	label_id
این چه محص	1
موبایل	0
لپ تاب	0
آبمیوه	1
هویج	1
پیاز	1
تلویزیون	0
کدو	1

و با استفاده از **train\_parsbert.py** دیتای خود را به مدل آموزش می دهیم و در انتها مدل را ذخیره میکنیم.

```
{'eval_loss': 0.5524888038635254, 'eval_runtime': 0.2039, 'eval_samples_per_second': 24.518, 'eval_steps_per_second': 24.518, 'epoch': 1.0}
{'eval_loss': 0.43550243973731995, 'eval_runtime': 0.1727, 'eval_samples_per_second': 28.954, 'eval_steps_per_second': 28.954, 'epoch': 2.0}
{'eval_loss': 0.3763268291950226, 'eval_runtime': 0.1885, 'eval_samples_per_second': 26.52, 'eval_steps_per_second': 26.52, 'epoch': 3.0}
{'train_runtime': 17.6266, 'train_samples_per_second': 1.362, 'train_steps_per_second': 1.362, 'train_loss': 0.45007352034250897, 'epoch': 3.0}
100% | 24/24 [00:17<00:00, 1.36it/s]
```

مشخصات مدل برای **fine-tune**

```
training_args = TrainingArguments(
    output_dir="./results",
    evaluation_strategy="epoch",
    learning_rate=2e-5,
    per_device_train_batch_size=1,
    per_device_eval_batch_size=1,
    num_train_epochs=3,
    weight_decay=0.01,
    logging_dir="./logs",
)
```

در انتها دسته بندی جنس / محصول مورد نظر پیشبینی می شود. اگر محصول مورد نظر به درستی پیش بینی شود. صفحه گوگل مربوطه به صورت خود کار درانتهای برنامه باز می شود.

```
is related to the category: Groceries
```

See results  
about



(كدو تقبل) Pumpkin  
Squash



(کدو سبز) Courgette  
Summer squash

Squash(کدو)

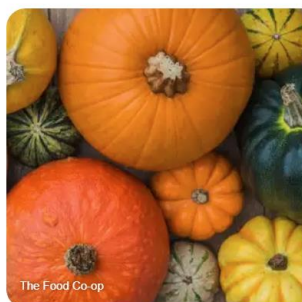
Plant :

## Overview

Lower classifications

### Benefits

## Types

Scient  
name

Cucurbita

Family

Cucurbitaceae



Wikipedia

<https://fa.wikipedia.org/wiki/کرو> · [Translate this page](#)

کدو - ویکی‌پدیا، دانشنامهٔ آزاد



Monica



Fast Model ▾

🗨️ Ask Monica

نکته: در صورت اجرای برنامه بر روی ترمینال پایتون حروف فارسی بهم ریختگی دارند.

با تشکر