8-28-2020

Intel

A Step in Reclaiming the PC Build Summit

Presented By: David Lee, Data Scientist, GA Data Consulting

Agenda

- Overview
- Problem Statement
- Process
- ML Models
- Conclusions
- Next Steps
- Q&A

Overview

- Workers are working remotely
- Manufacturing has slowed
- PC parts are in demand
- Delays in 7nm chips until 2022

By the time Intel finally launches its 7nm chips in 2022 and 2023, TSMC will likely have advanced to 3nm chips. Therefore, if AMD keeps designing smaller chips and simply places its faith in TSMC, it could eventually replace Intel as the "best in breed" CPU maker for the PC and data center markets.

~The Motley Fool~

Problem Statement

1. How can we predict and market the computer parts consumers will buy, better than our competitors?

Process

- Using Reddit posts as a tool for gaining insights
- Web Scraping and Data Cleaning

Web Scraping

110K

AMD Subreddit Posts

WebScraping

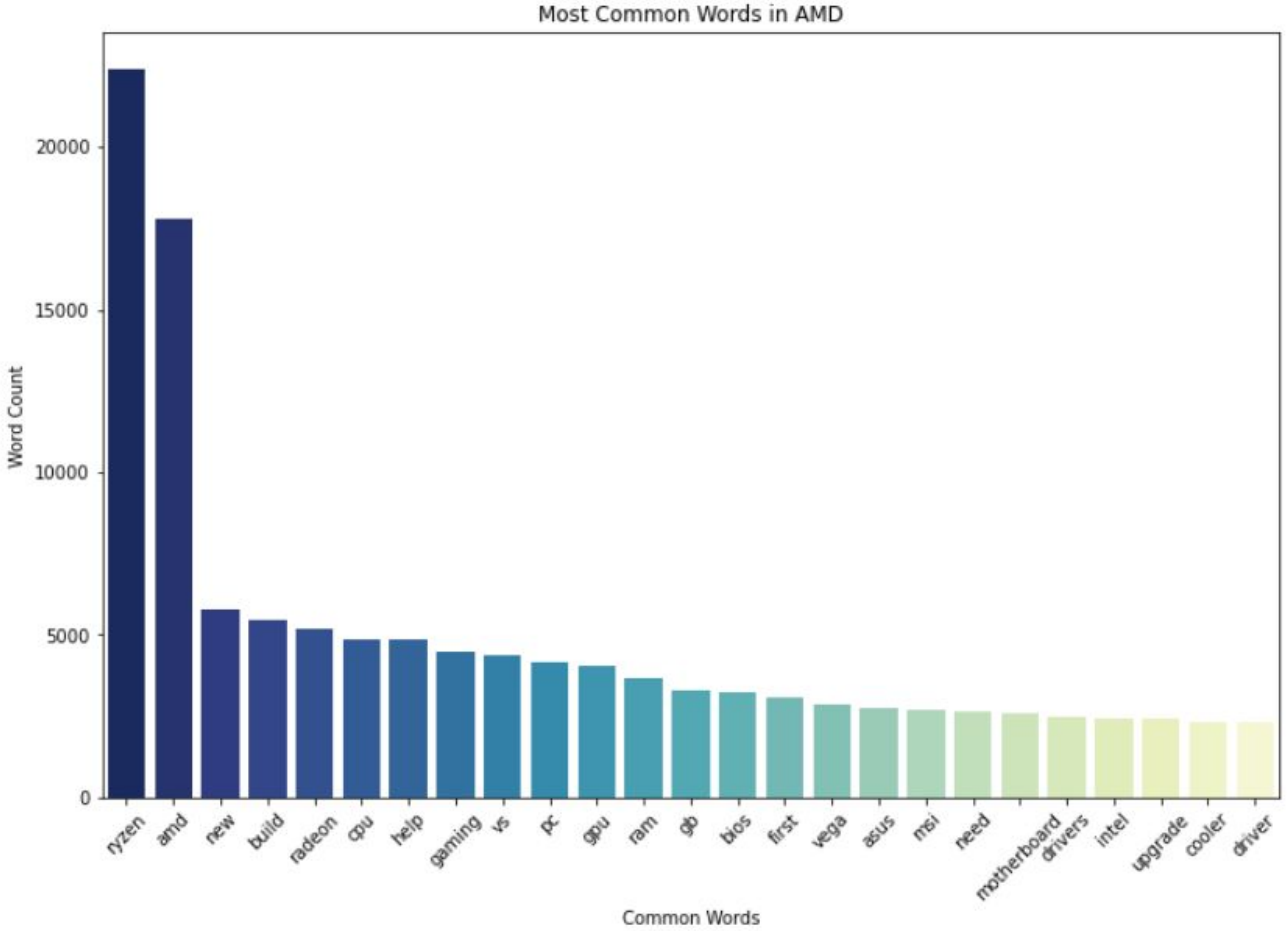
110K

Build a Pc Subreddit Posts

Posts with Content

78.56%

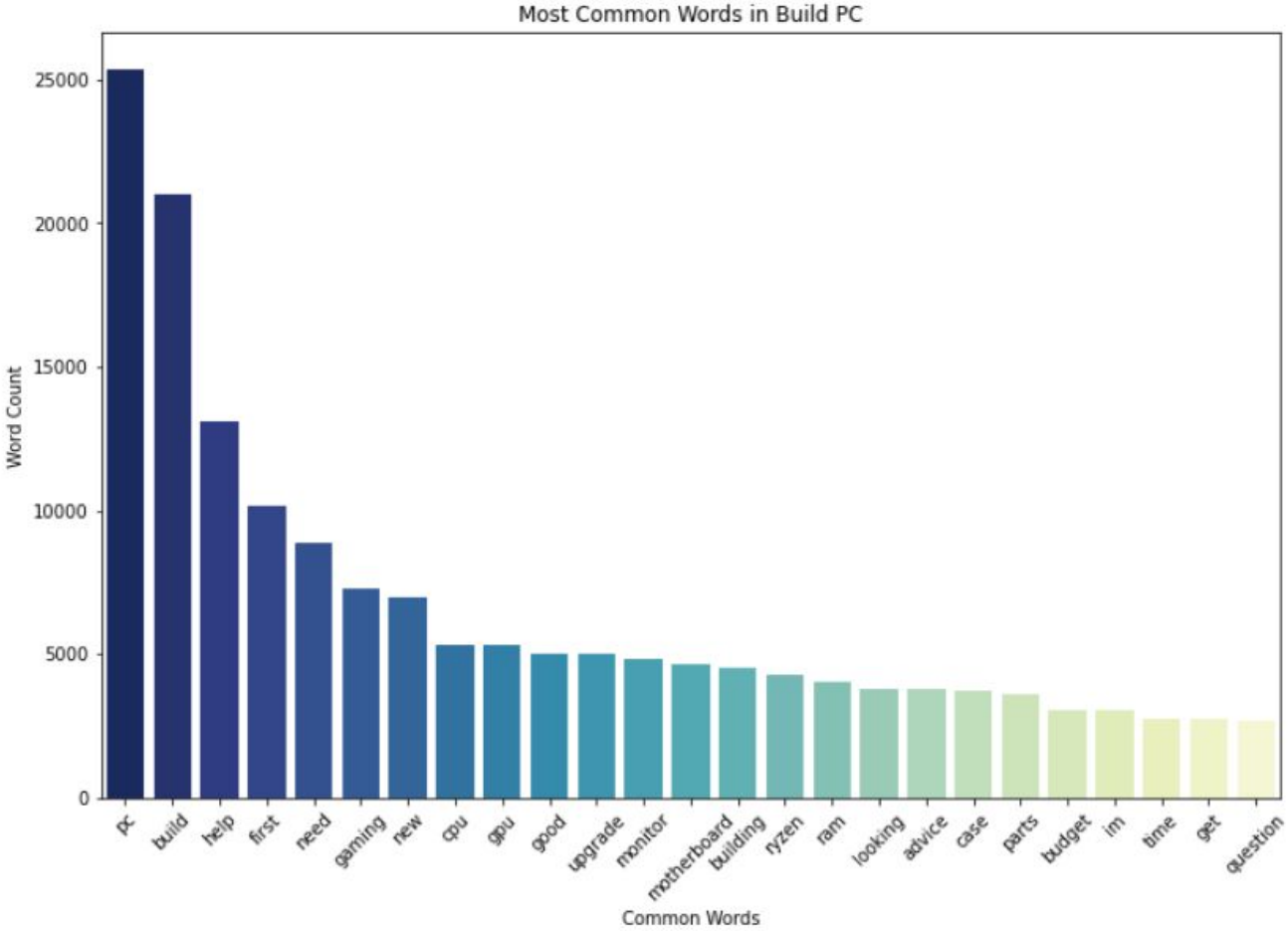
Posts containing content



Top 25 Words in AMD

Unique Words Not in Build a PC

- AMD
- ASUS
- Bios
- Cooler
- Driver
- Drivers
- GB
- Intel
- MSI
- Radeon
- Vega
- VS



Top 25 Words in Build a PC

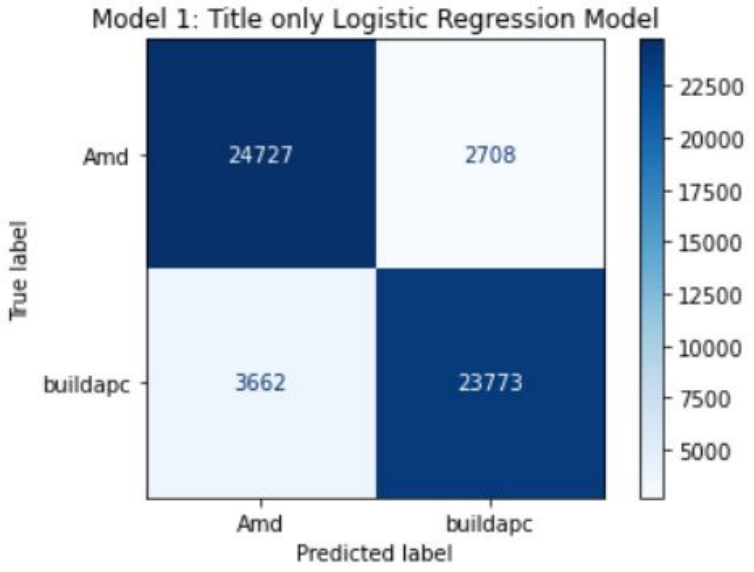
Unique Words Not in AMD

- Advice
- Budget
- Building
- Case
- Get
- Good
- Im
- Looking
- Monitor
- Parts
- Question
- Time

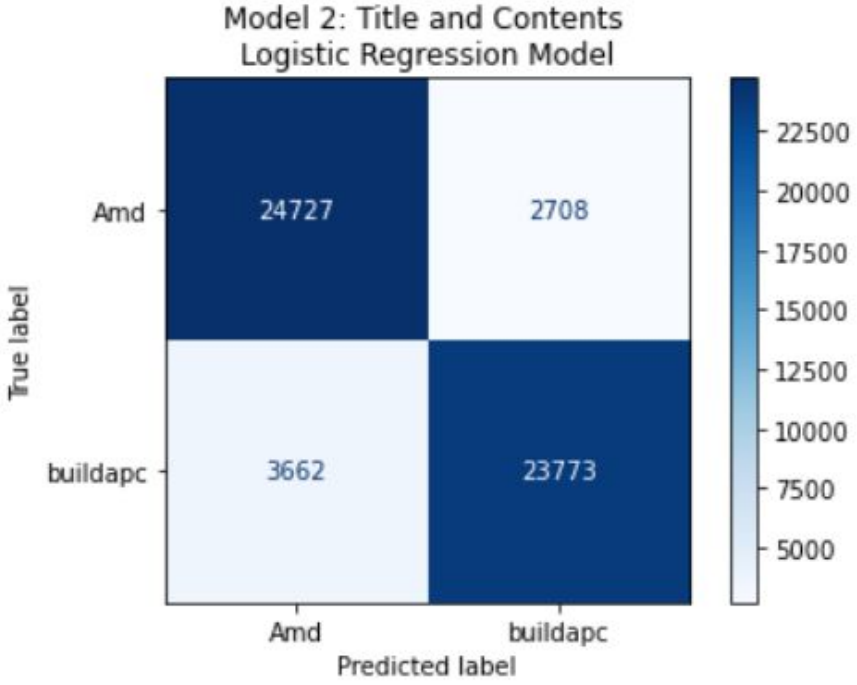
NLP Modeling

- Logistic Regression
- Naive Bayes
- Optimal Parameter Searching Modeling (Gridsearch)

Target audience

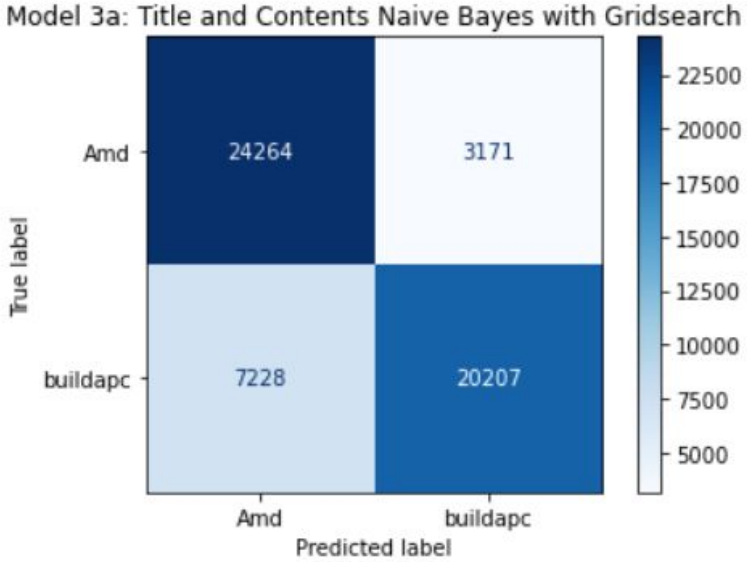


The Accuracy score is: 84.5%
The Missclassification rate is: 15.5%
The Sensitivity is: 85.7%
The Specificity is: 83.3%
The Precision is 83.7%

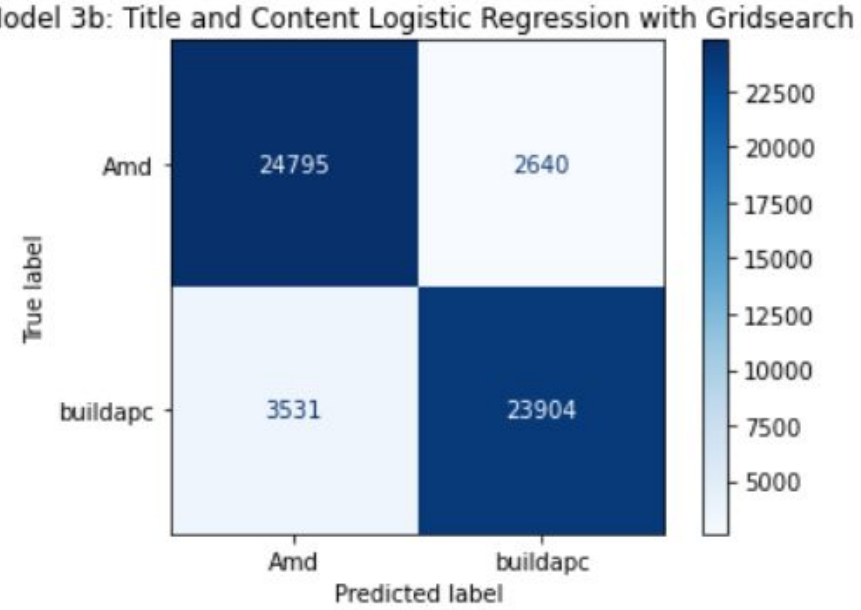


The Accuracy score is: 88.39%
The Missclassification rate is: 11.61%
The Sensitivity is: 86.65%
The Specificity is: 90.13%
The Precision is 89.8%

Target audience



The Accuracy score is: 81.0%
The Missclassification rate is: 19.0%
The Sensitivity is: 73.7%
The Specificity is: 88.4%
The Precision is 86.4%



The Accuracy score is: 88.8%
The Missclassification rate is: 11.25%
The Sensitivity is: 87.1%
The Specificity is: 90.4%
The Precision is 90.05%

Conclusions:

With a high accuracy model, we believe we can build a model that provides valuable market insights on multiple social platforms, on all of your competitors.

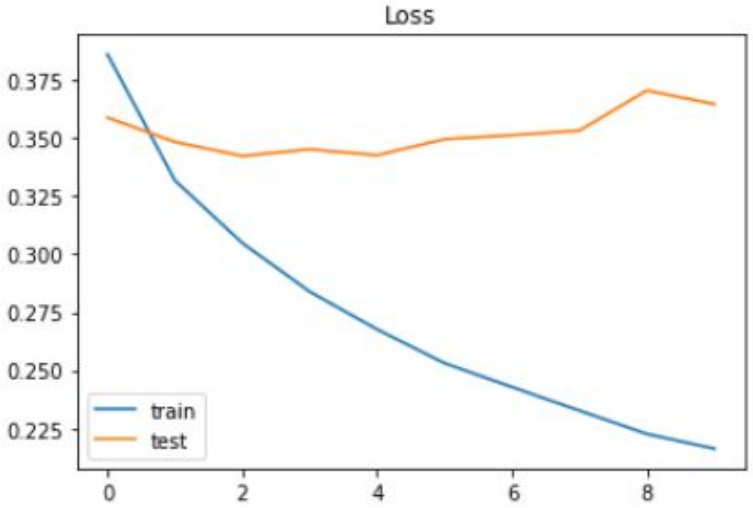
Next Steps:

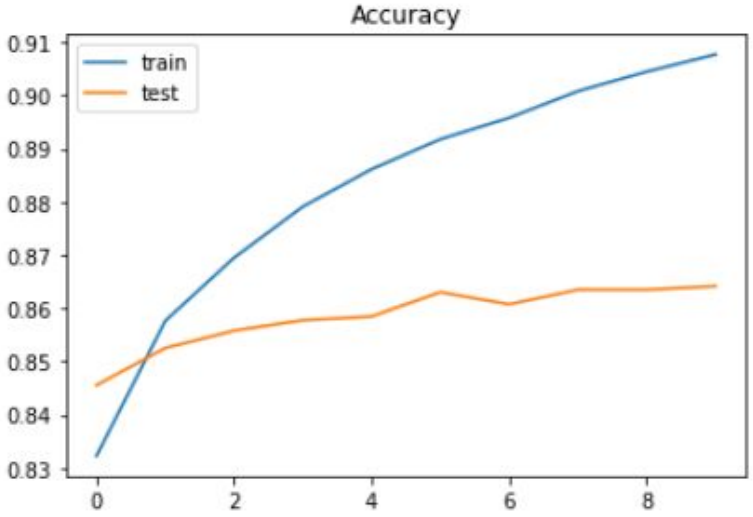
- Building an auto updating production model

- Scraping data on many social platforms, on all competition.

- Creating a market and development strategy that is never late

LSTM RNN Model on Title Only






```
686/686 [==============================] - 35s 51ms/step - loss: 0.3618 - accuracy: 0.8651
Test set
  Loss: 0.362
  Accuracy: 0.865
```

LSTM RNN Model Test on Title Only

Video Radeon Driver 20.8.3 vs 20.1.4 - 20.8.2 in 7 games (RX 570) - With Microsoft Flight Simulator 2020

youtu.be/6M5Ajw...


```
# An AMD Post
new_post = ['Radeon Driver 20.8.3 vs
seq = tokenizer.texts_to_sequences(ne
padded = pad_sequences(seq, maxlen=MA
pred = model.predict(padded)
labels = ['AMD', 'Build a PC']
print(pred, labels[np.argmax(pred)])
```

```
Last executed at 2020-08-28 10:21:39 in 208
[[0.97912174 0.02087824]] AMD
```

Posted by u/[redacted] 15 hours ago

Miscellaneous **Im a bit overwhelmed with the many different graphics cards. Please help.**

Hi guys, im currently looking to upgrade with a new CPU, GPU and get a bit more RAM. So i am currently trying to pick out a new GPU and im not ultra knowledgeable about building PCs or picking component but i decided to get a graphics card from the NVidia RTX 2080 line. Now i have checked several websites that offer PC components and i am a bit confused about how many graphics cards with seemingly the same name, the same (or similar) specs etc are on sale for wildly different prices (even on their official website). Now how do i actually find out the differences in order to choose which one i might

```
# A Build a PC Post
new_post = ['Im a bit overwhelmed with the ma
seq = tokenizer.texts_to_sequences(new_post)
padded = pad_sequences(seq, maxlen=MAX_SEQUEN
pred = model.predict(padded)
labels = ['AMD', 'Build a PC']
print(pred, labels[np.argmax(pred)])
```

```
Last executed at 2020-08-28 10:21:42 in 60ms
[[9.6334820e-04 9.9903667e-01]] Build a PC
```

Q&A and Many Thanks