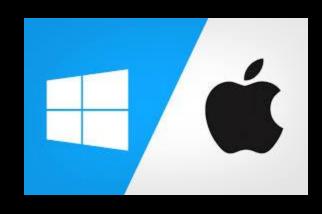
# Webscraping & Classification Model - PC vs Mac Gaming



February 2020

### **Agenda**

- Problem Statement
- Key Objectives
- State of the Market
- Classification Modeling & Evaluation
- Recommendations & Next Steps

### **Problem Statement & Key Objectives**

#### **Context:**

For too long, PC's dominance in the Gaming, Enthusiast, and eSports space has gone unchecked.

Tim Cook, CEO of Apple is seeking to explore new market opportunities in order to reverse stagnant Apple unit sales. Therefore Apple has recently hired a new executive to spearhead this initiative, named *Cim Took*.

Mr Took has scheduled time with you to discuss initial steps and analyses he will need. In order to prepare for this meeting you turn to your favorite source of consumer sentiment: *Reddit* 

#### Subreddits evaluated:

- PC Gaming Master Race
- MacGaming

#### **Key Objectives:**

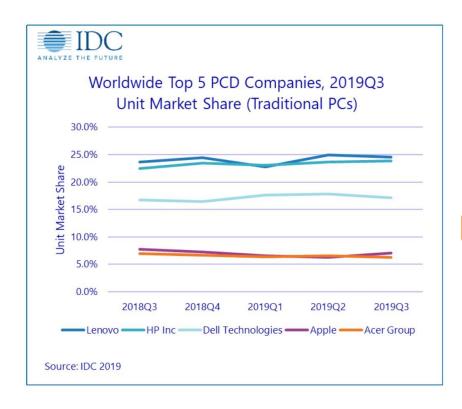
- PULSE OF THE CUSTOMER: Begin prototyping a Classification System for ongoing analysis of platform-agnostic gaming consumer comments. Eventually this will expand to include sentiment analysis across various platforms (not just reddit) and customer targeting for ad campaigns
- Evaluate model how accurately can it identify PC Gaming apart from Apple Gaming

### **Executive Summary**

- Key objectives: Prototype a Classification System for ongoing analysis of gaming consumer comments to track performance of Apple Gaming initiative. Eventually model will expand to include sentiment analysis across various platforms (not just reddit) and customer targeting for ad campaigns
- PC Gaming Hardware Market is a large opportunity estimated to be over \$23 Billion in 2017, shows strong growth and is dominated by PCs. Apple does not currently have strong performing or price-effective solutions for this space
- PCs also dominate the Reddit gaming community as well as high visibility (60 million+ viewer) eSports events
- Over 3.5 Million records were scraped from Reddit however after duplicates and balancing classes 2100 per class were kept
- 7 distinct machine learning models were applied and SVC exhibited the highest accuracy and strongest ability to distinguish between PC & Apple Gaming with an R<sup>2</sup> score of 0.8980 and ROC Curve with AUC 0.947 (showing strong performance with Sensitivity and Specificity)
- Model Coefficients & Insights suggest gamers prefer custom builds with RGB, high performance and SSDs
- Recommendations:
  - Customer Survey to corroborate insights on customer gaming needs (e.g. FPS, CUDA cores, compatibility issues) to influence roadmap
  - Continue building out robustness of classification model: additional feature reduction/refinement, additional feature engineering, lemmatization, experiment with different platforms such as Amazon reviews, Newegg reviews, Twitter
  - Begin cultivating esports partnerships and cyberathlete sponsorships

# Background & State of the Market

# State of the Market: Apple 19Q3 Unit Shipments & Unit Share Declined Compared To 18Q3



Top Companies, Worldwide Traditional PC Shipments, Market Share and Year-Over-Year Growth, Third Quarter 2019 (Preliminary results, shipments are in millions of units)							
Company	3Q19 Shipments	3Q19 Market Share	3Q18 Shipments	3Q18 Market Share	3Q19/3Q18 Growth		
1. Lenovo	17.3	24.6%	16.2	23.6%	7.1%		
2. HP Inc.	16.8	23.8%	15.4	22.5%	9.3%		

23.8%	15.4	22.5%	9.3%	
17.1%	11.5	16.8%	5.3%	
7.1%	5.3	7.7%	-6.1%	
	3.3	,.,,	0.170	
6.3%	4.8	7.0%	-7.2%	
	17.1%	17.1% 11.5	17.1% 11.5 16.8%	17.1% 11.5 16.8% 5.3%

Source: IDC Quarterly Personal Computing Device Tracker, October 10, 2019

# PCs Dominate High-Growth eSports Events And Corresponding Products

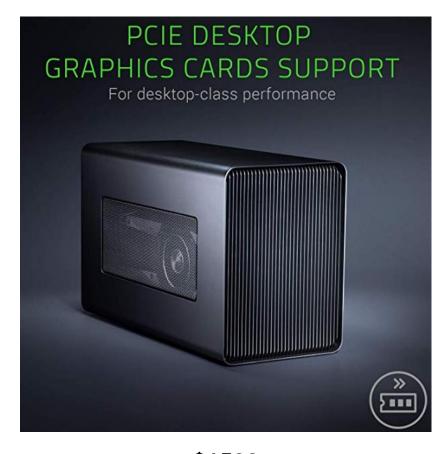


**100 MILLION unique viewers** watched the 2019 League of Legends tournament which had a \$2.5M grand prize (that's more than the Super Bowl)

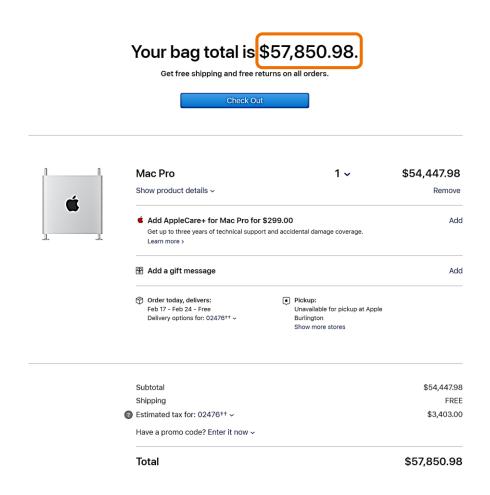


- 2.5 BILLION Video Gamers all over the world
- 62% of customers prefer playing games on PC / Mac
- Total PC Gaming Hardware Industry Revenue in 2017 was \$23 Billion
- Steam game sales revenue reached \$4.3B in 2017, growing 19% YoY

### State of the Market: Apple Gaming

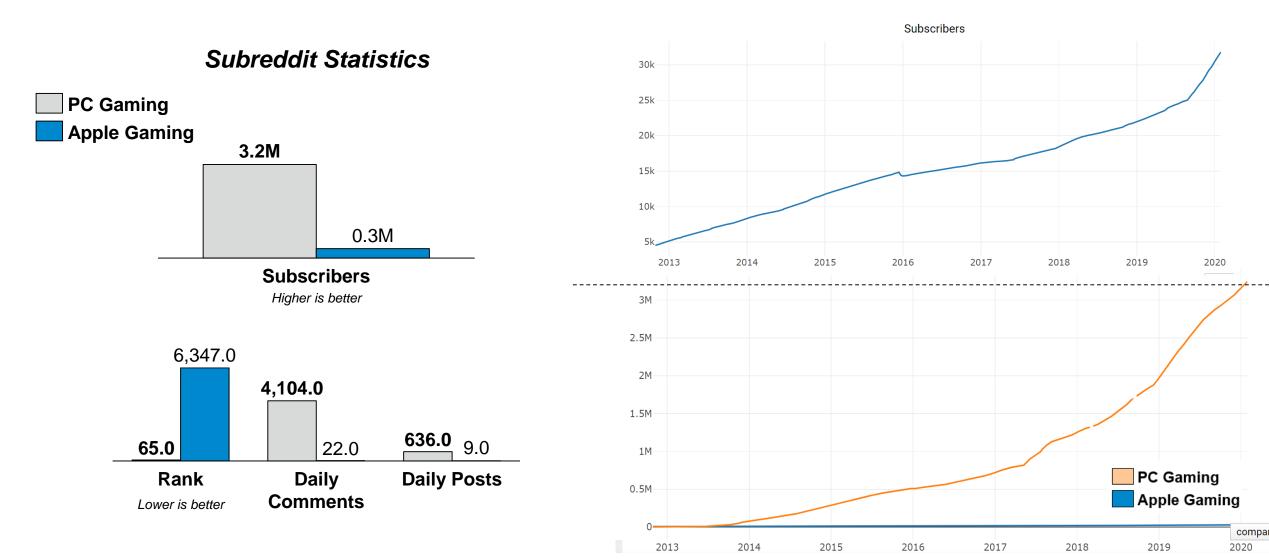


\$1500 Toaster \$299 Just for the enclosure. \$1200 for a topline GPU



Note: \$3,400 in state sales tax

### **PCs Also Dominate Reddit Gaming Community**



Source: https://subredditstats.com/r/pcmasterrace; https://subredditstats.com/r/macgaming

# Modeling

#### **Model Overview**

#### Reddit **Data Collection**

Cleaning,

Processing,

Vectorization

collected

TfidfVectorizer

Deduplication

PushShift API

Over 1800 days

- English stopwords included
- English language standardization

• 3 million+ records (5GB) were

Removed irrelevant or international words/features

**Modeling & Hypertuning** 

- Support Vector Machines 89.804% test accuracy (best model)
- Logistic Regression
- XGBoost Classification

Gaussian Naïve Bayes

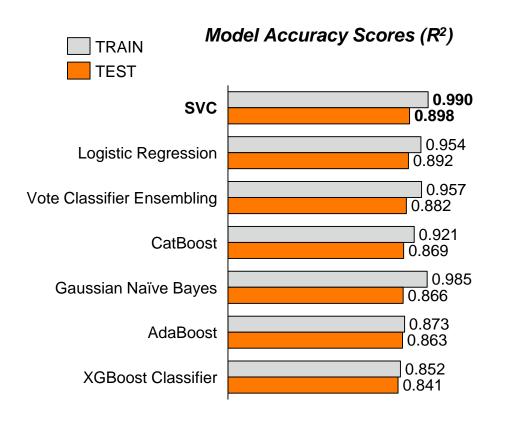
Apple only 2000 uniques

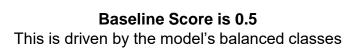
PC was reduced to 2000

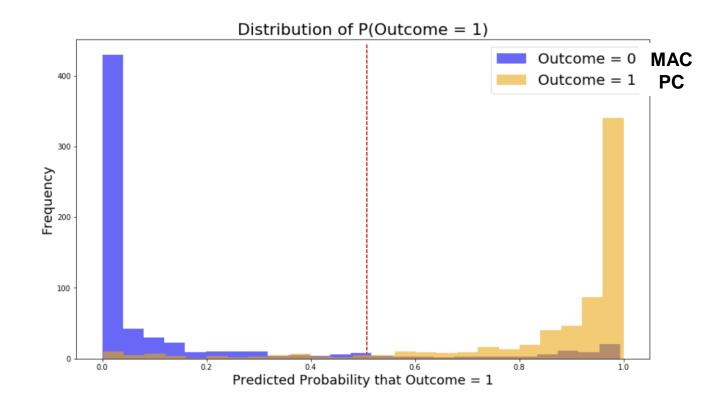
· In order to maintain balanced classes,

- AdaBoost
- CatBoost
- Voting Classifier Ensembling

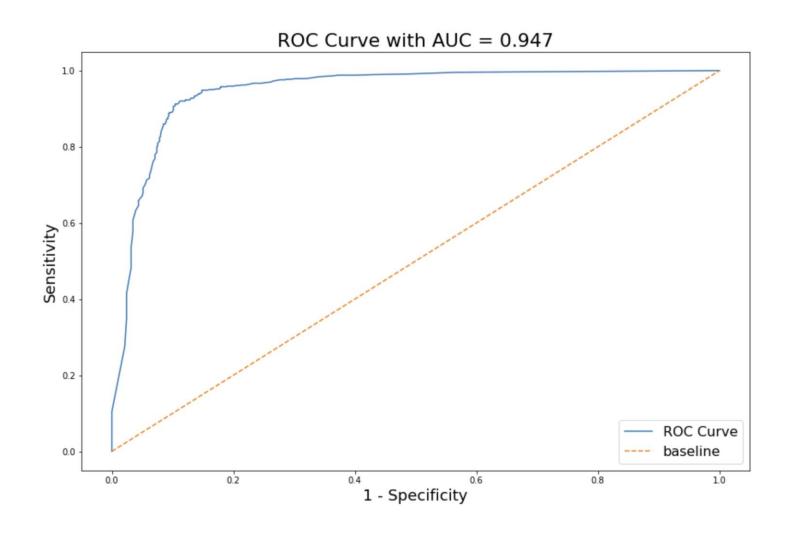
# Model Evaluation: High Accuracy & Strong Ability To Distinguish Between Classes (Mac vs PC)







#### **Model Evaluation: Effective Distinction Between Classes**



#### 0.932

 Sensitivity – how well the model classifies PC Gaming

#### 0.864

 Specificity – how well the model classifies Mac Gaming

## Insights: Leading Coefficients Suggest Gamers Prefer Custom Builds With RGB Lights, High Performance And SSDs

Leading words that had the highest coefficients (absolute values):

- **Dead** researching own build after dead Apple product
- Questions on custom builds
- GPU CPU seeking methods of improving performance
- Finally complete indicates custom PC build
- **Z370** intel chipset
- Rgb strips apple has no RGB
- **Xbox** games seeking better performance/gameplay experience than console
- New ssd
- Corsair

- Halo
- Underperforming Apple customers
- Arrived today components arriving for custom PC builds
- Final build
- Egpu Apple install base researching gameplay improvement options
- Video card
- Low gaming settings Apple
- Feels good customer satisfied with PC custom builds

### **Recommendations & Next Steps**

- Survey to bring out insights on customer gaming needs (e.g. FPS, Cuda Cores, compatibility issues) to influence roadmap
- Customer segments to include:
  - Existing PC Gaming enthusiasts
  - Existing Apple fanboys
  - Esports athletes
  - Regular consumers (gauge gaming appeal)
- Begin working on eSports partnerships and cyberathlete sponsorships
- Continue building out robustness of classification model: additional feature reduction/refinement, additional feature engineering, lemmatization, experiment with different platforms such as Amazon reviews, Newegg reviews, Twitter
- Apple to further evaluate addressing custom-builds either through easier installation of operating system or partnerships with component manufacturers (e.g. hardware + MacOS bundling), and RGBs...
- Build financial multi-year projections to estimate Apple entry and cannibalization of PC Gaming market