

# PC Build Value Performance Application

Alexander Kim

# Proposal

- Create an application in Rshiny for users to analyze their PC build's value over performance using supervised regression
- Compare current PC builds from user-submitted parts on pcpartpicker, and analyze the price per performance
  - Gradient Boosting (XGBoost / LightGBM) or Random Forest on features: price, benchmark, power, memory size...
- Train to predict synthetic score per dollar
- Why: Answers “how good is this?” estimate and can then rank by predicted value

# Data

- Each computer part will be web scrapped from pcpartpicker (<https://pcpartpicker.com/builds/>) (Currently blocked)
  - Provides a more detailed description of specifications for each part
    - Core count, Thread count, L2 cache, TDP (watt), Clock Speed, Memory
- Benchmarked performance of cpu/gpu curated from
  - <https://gpu.userbenchmark.com/>
  - <https://www.cpubenchmark.net/>
- Uses Selenium WebDriver to render JavaScript content.
  - Waits 5 seconds to allow JavaScript content to fully load
- BeautifulSoup: To parse the HTML
- Searches for all elements where the id matches
  - Extracts:
    - CPU Name (<span class='prcname'>)
    - CPU Mark / Score (<span class='count'>)
    - Price (<span class='price-neww'>)
- Converts the data list into a pandas DataFrame

```
1 CPU Name,CPU Mark,Price
2 Intel Core i5-11300H @ 3.10GHz,"10,875",$309.00*
3 Intel Core i5-1335UE,"10,863",$312.00*
4 Intel Xeon E5-4650 v3 @ 2.10GHz,"10,838","$3,838.00*"
5 AMD Ryzen 3 PRO 4200G,"10,831",NA
6 AMD Ryzen 5 4500U,"10,825",NA
7 Intel Xeon E5-4650 v2 @ 2.40GHz,"10,818",$249.95
8 Mediatek MT6989,"10,792",NA
9 Intel Core i7-6800K @ 3.40GHz,"10,734",NA
10 Intel Core i7-9750H @ 2.60GHz,"10,721",$960.00*
11 Intel Xeon E5-2470 v2 @ 2.40GHz,"10,692",$49.95*
12 Intel Core i5-8600K @ 3.70GHz,"10,667",NA
```

# EDA

- Most CPUs fall between 3500 and 10000 benchmark points
  - Suggests the mid-range market has broad performance
- Price does not equate performance as much in the <\$1000 group
  - Mid-range market can't assume higher price always means better performance
  - There are more “value” CPUs that offer good performance at lower price
- CPU mark and Price correlation is .3 - weak to moderate positive correlation
  - Price alone is not a reliable predictor of performance

