Customer clustering on RV1000/RV1000A users

Customer clustering allow us to understand the patterns that differentiate our customers. Here are some ideas of what we might do with clustering analysis. Increased understanding of customer usage behavior. Develop products which best appeal to different segments of customers. Build loyal relationships.

Data

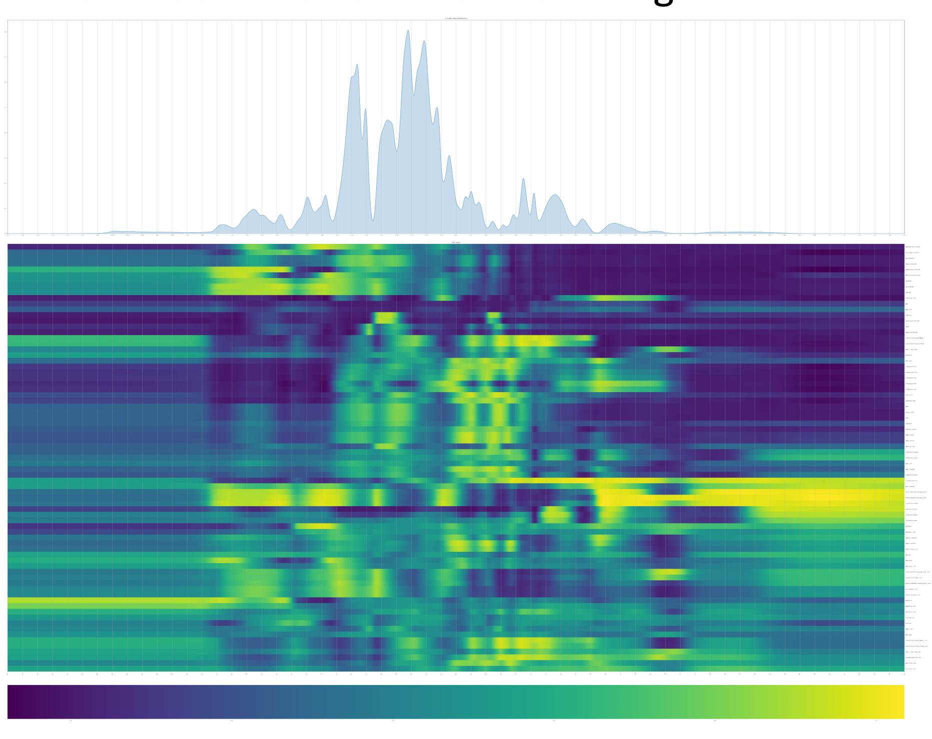
- Time: From 2019-08 till 2020-04.
- 166,684 unique robots.
- 990,478 data points(Each data point is the biweekly usage info of a robot)
- 119 features

Models

Deep learning – Variational Autoencoders (VAEs)

Model 1

- Dimension reduction method: For a given set of possible encoders and decoders, we are looking for the pair that keeps the maximum of information when encoding and, so, has the minimum of reconstruction error when decoding.

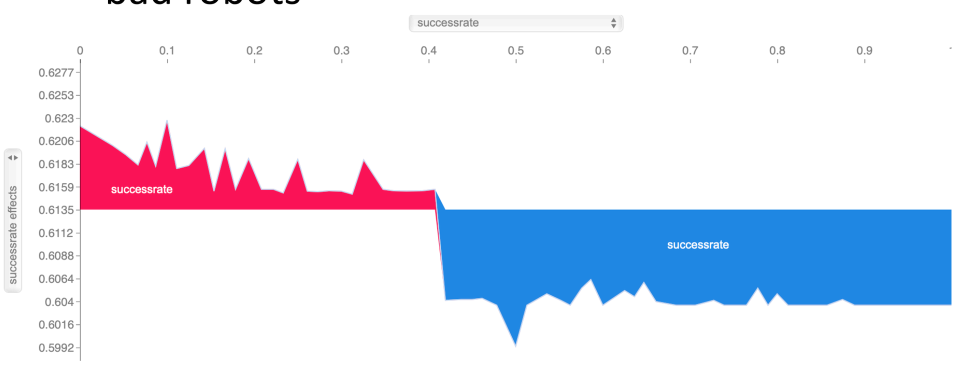


- Cluster of very bad robots: 13%
- Cluster of relatively bad robots: 33%

Machine learning– K means clustering

Model 2

- K means algorithm is an iterative algorithm that tries to partition the dataset into K pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group.
  - Cluster of new robots: 24%
  - Cluster of good robots: 31%
  - Cluster of bad robots: 45%
- **Thresholds of being bad robots:**
  1. Success rate > 0.45: less likely to become bad robots

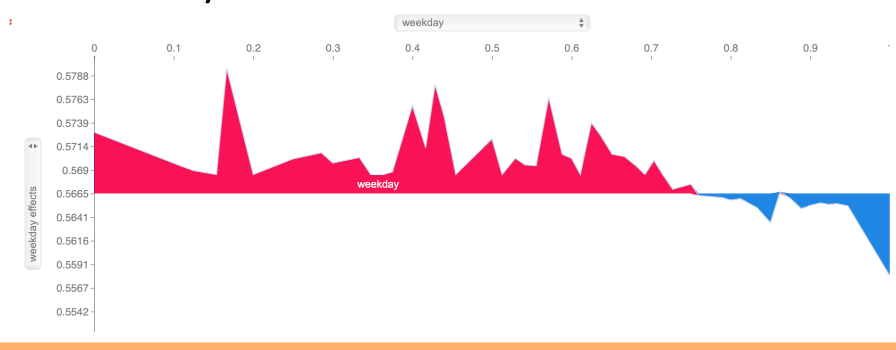


  2. Percentage of runs ending in contacted charging dock status > 80%: less likely to become bad robots

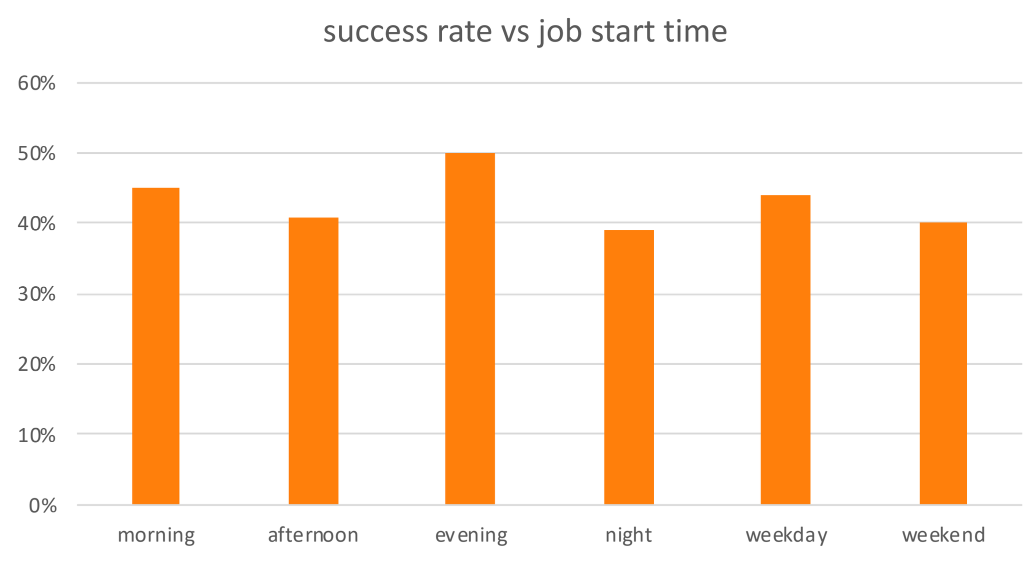
Combination of Variational Autoencoders & K means clustering

- **Mode 3**
  - Cluster of good, old, success on weekend customers: 22%
  - Cluster of good, bad morning run customers: 13%
  - Cluster of good, success on weekday and afternoon customers: 4%
  - Cluster of new robots: 11%
  - Cluster of good, old, success in the morning customers: 10%
  - Cluster of bad, failed in the morning customers: 30%
  - Cluster of new, bad, morning run customers: 10%
- **Threshold of being bad:**
  - Avg number of errors received per job > 0.3: more likely to become bad robots

- **Mode 4**
  - Cluster of good, weekend run customers: 32%
  - Cluster of good customers: 13%
  - Cluster of good, weekday run, success in the morning customers: 12%
  - Cluster of bad robots: 24%
  - Cluster of bad, new robots: 19%
- **Threshold of being bad:**
  - Percentage of weekday runs > 0.75: less likely to become bad robots

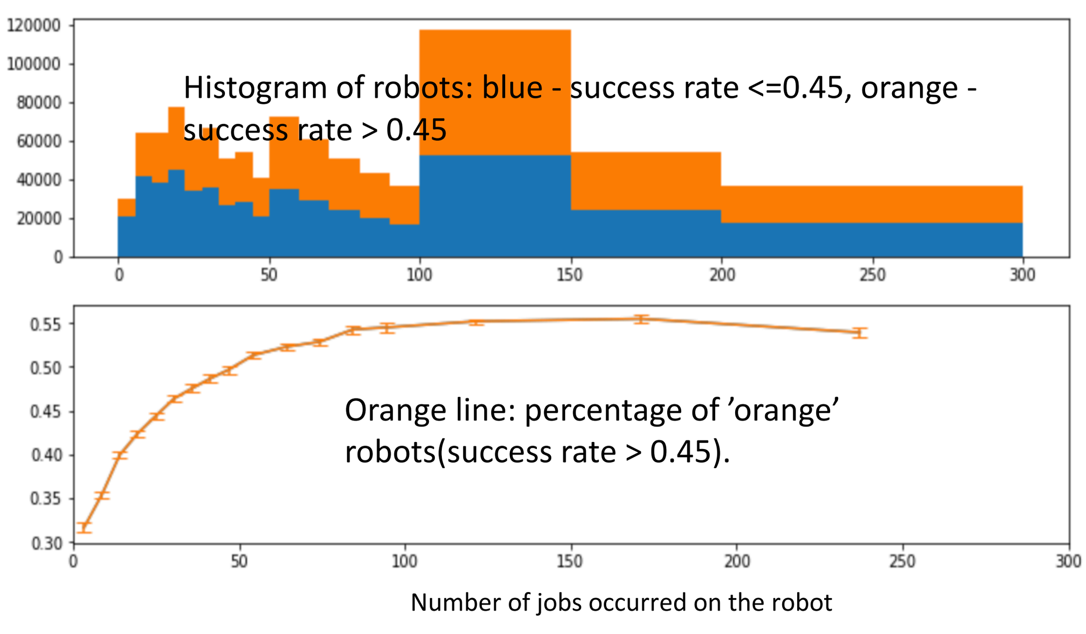


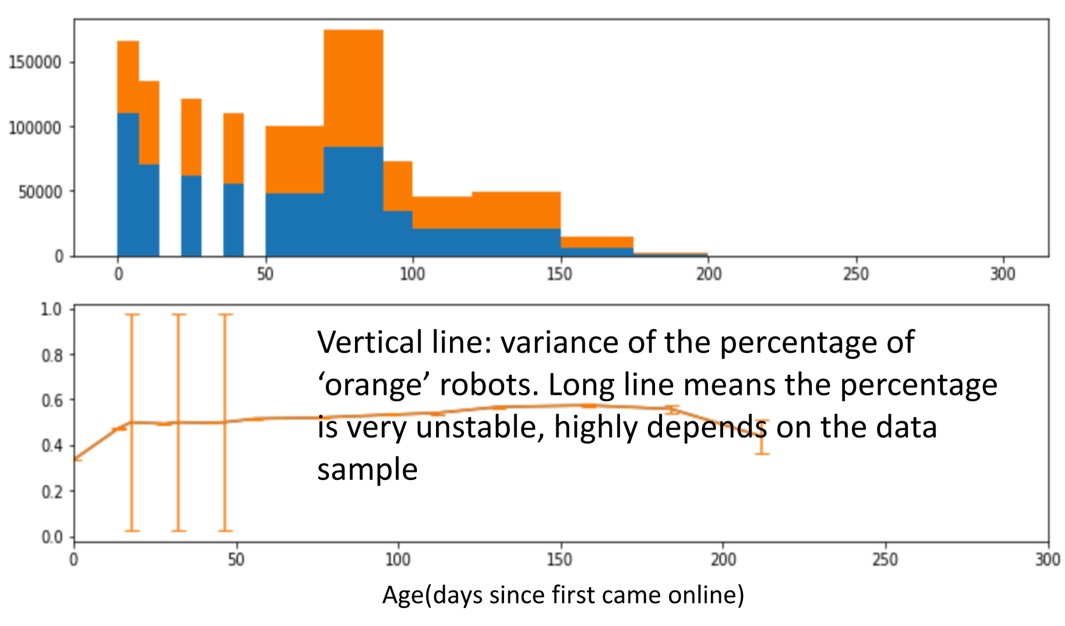
Analysis of usage info



Jobs happened in the morning, evening and on weekday have larger success rate

Using threshold(success rate > 0.45) found before to separate robots into successful and failed robots

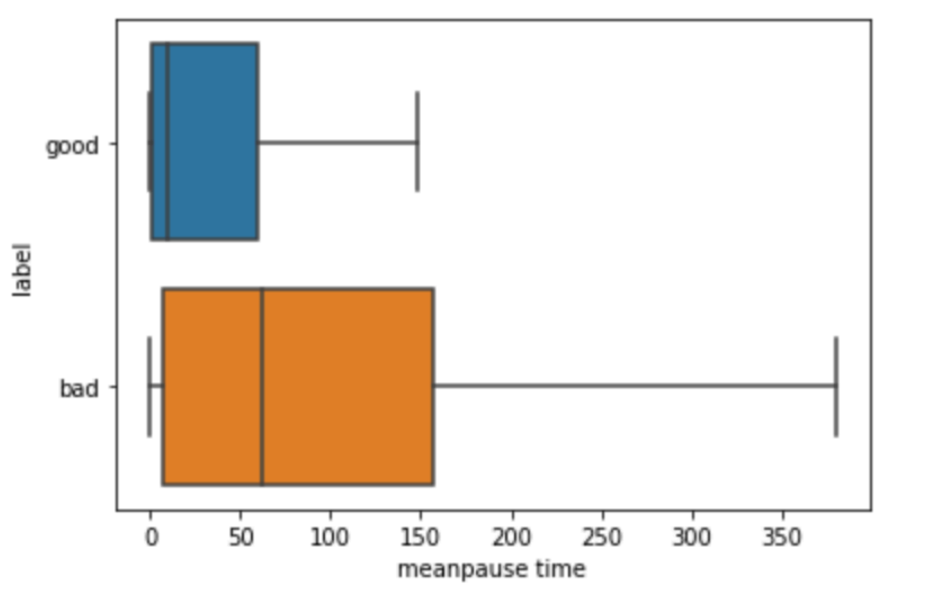




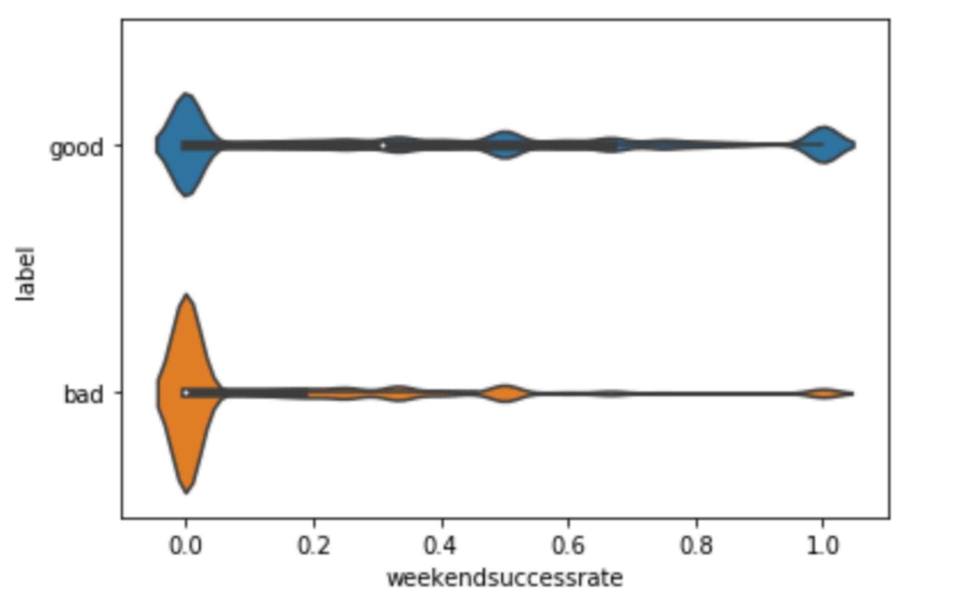
- 0~80 jobs: the probability of being successful continues to increase with increasing number of jobs
- More than 80 jobs: the probability of being successful becomes stable and is close to 0.55
- 0~50 days: the probability of being successful increases with increasing age and is very unstable
- More than 50 days: the probability of being successful becomes stable and is close to 0.55

Analysis of the final model result

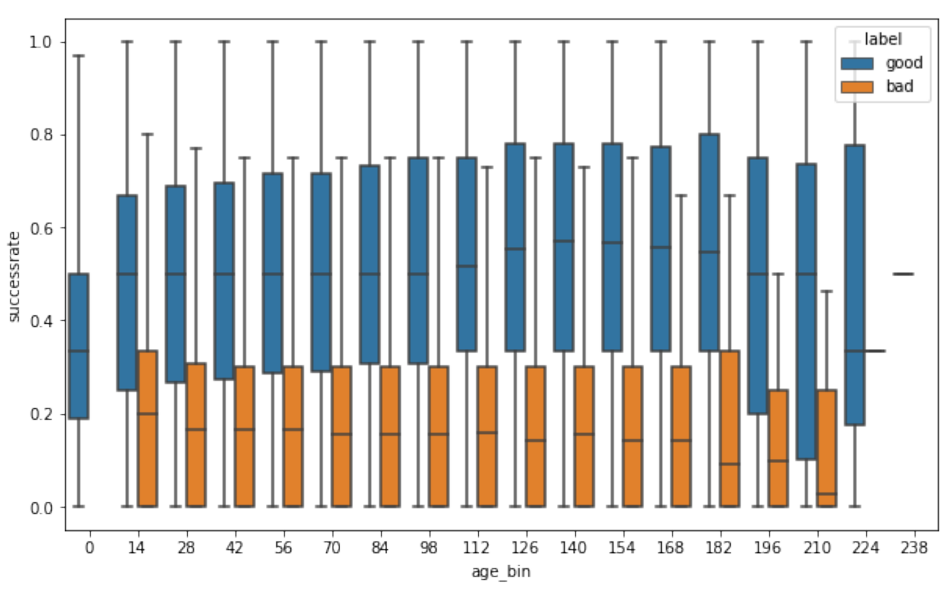
Combine cluster results of 4 models above to generate the final results.
Bad robots: 12% Good robots: 88%



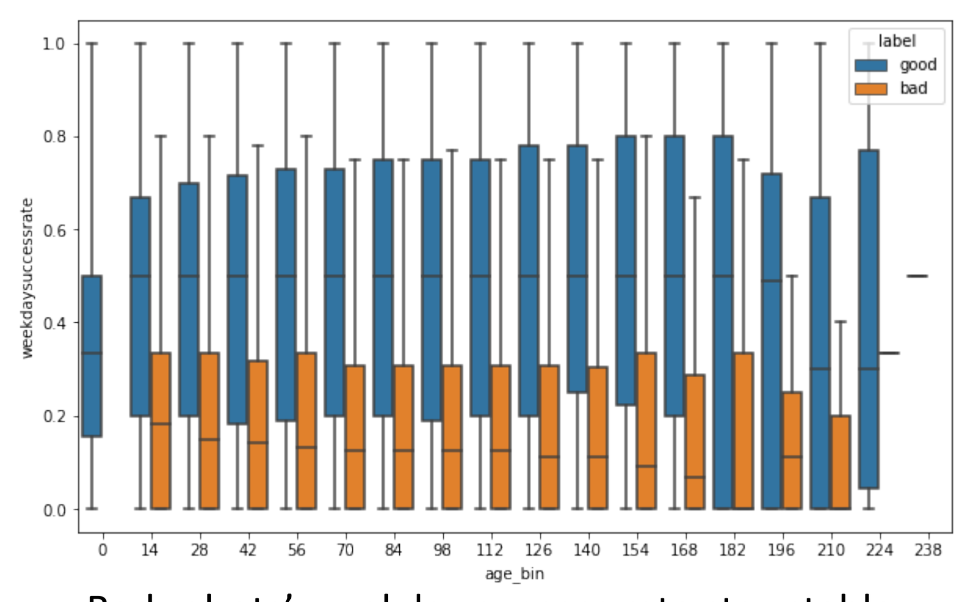
- Bad robots have larger avg pause duration(minutes).
- Median pause time of bad robots is around 60 mins



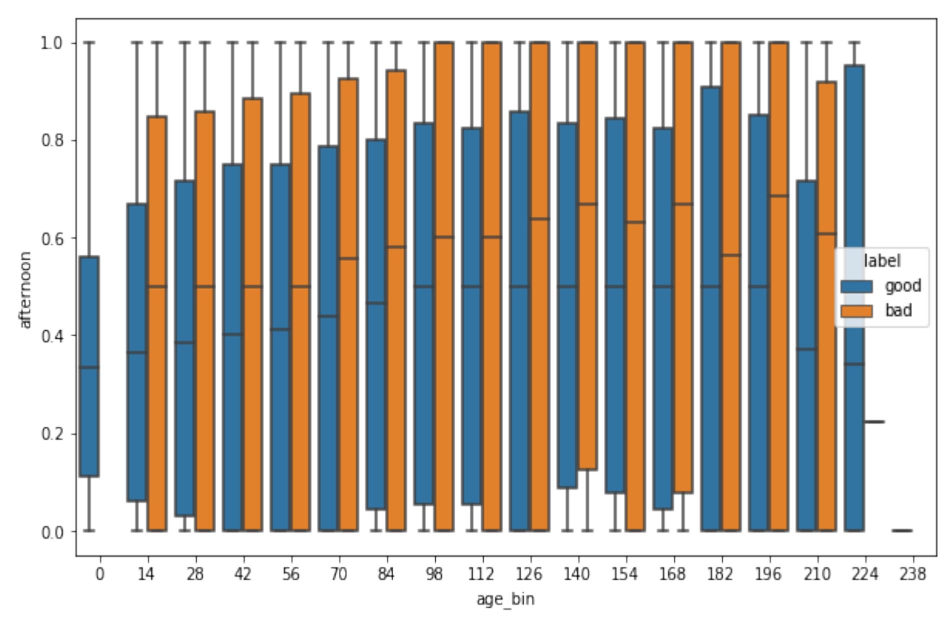
- Good robots have larger weekend success rate.
- Median weekend success rate of bad robots is around 0



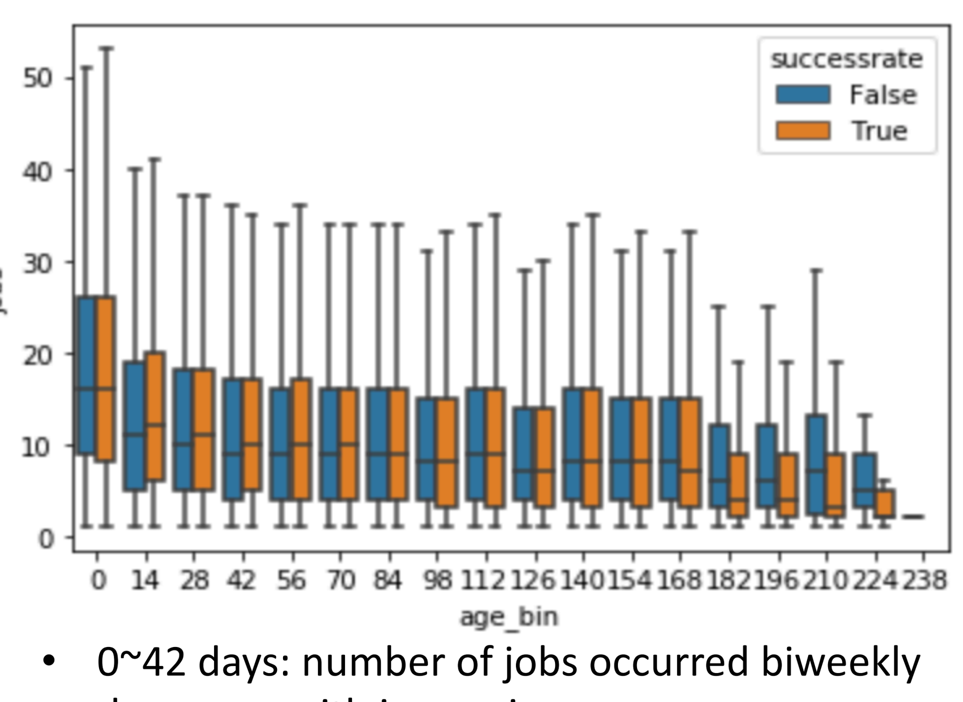
- Bad robots' success rate stay stable when getting old
- Good robots' success rate increase as getting old
- The success rate difference between good and bad robots becomes larger and larger



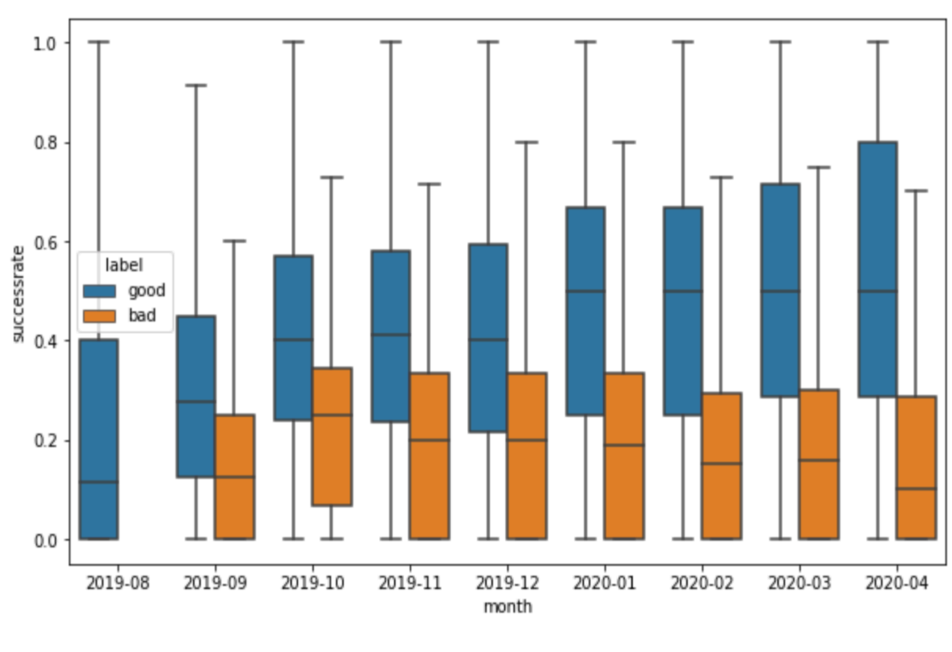
- Bad robots' weekday success rate stay stable
- Good robots' weekday success rate increase
- The weekday success rate difference between good and bad robots increases



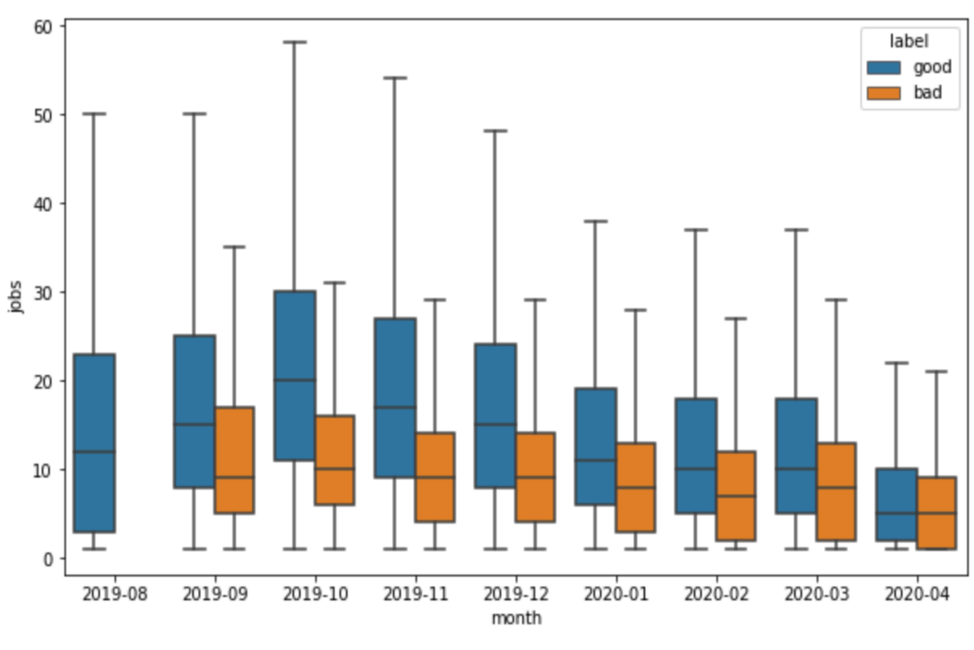
- Robots have more afternoon runs when getting old



- 0~42 days: number of jobs occurred biweekly decreases with increasing age
- After 42 days: number of jobs becomes stable



- The success rate difference between good and bad robots becomes larger and larger when approaching 2020-04



- Number of jobs in 2 weeks for bad robots is smaller
- Number of jobs for good robots decreases