To Bid, or Not to Bid: That is the Question

Rachel Weber

A Classification Problem

Question to address:

When should an organization bid at auction?

Knowing whether or not an organization is likely to win at auction requires knowledge of previous performance

- Winning is influenced by a multitude of factors, some clean-cut, others more obtuse
- Being able to take all variables into account when deciding when to bid allows for reduced risk in bidding and efficient use of advertising dollars

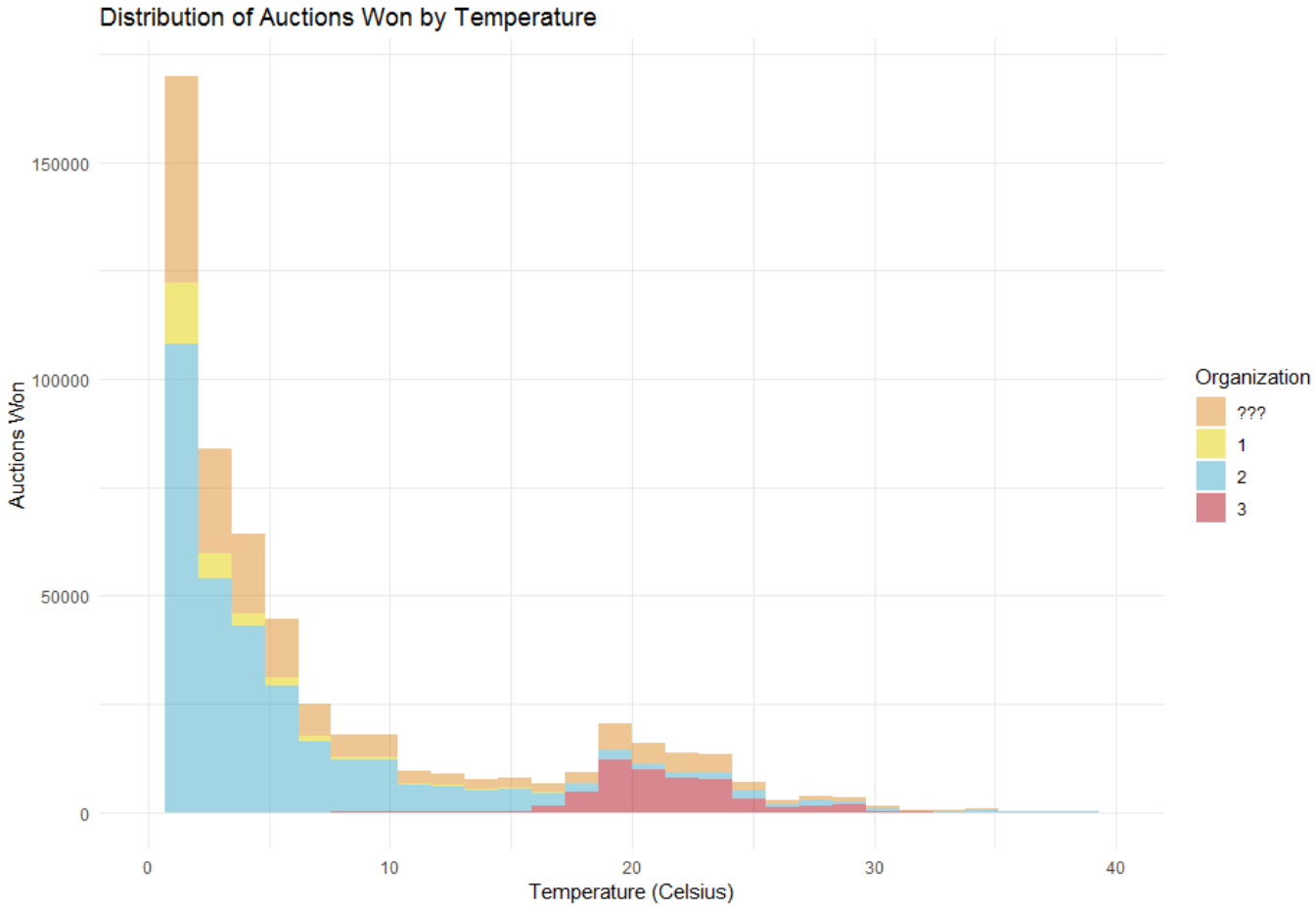
The Solution: K Nearest Neighbors

Why Use?

- It's a common, easy to interpret, algorithm used in classification problems
- Can handle large amounts of data
- Uses testing and training datasets to build a model and assess its performance

Implementation

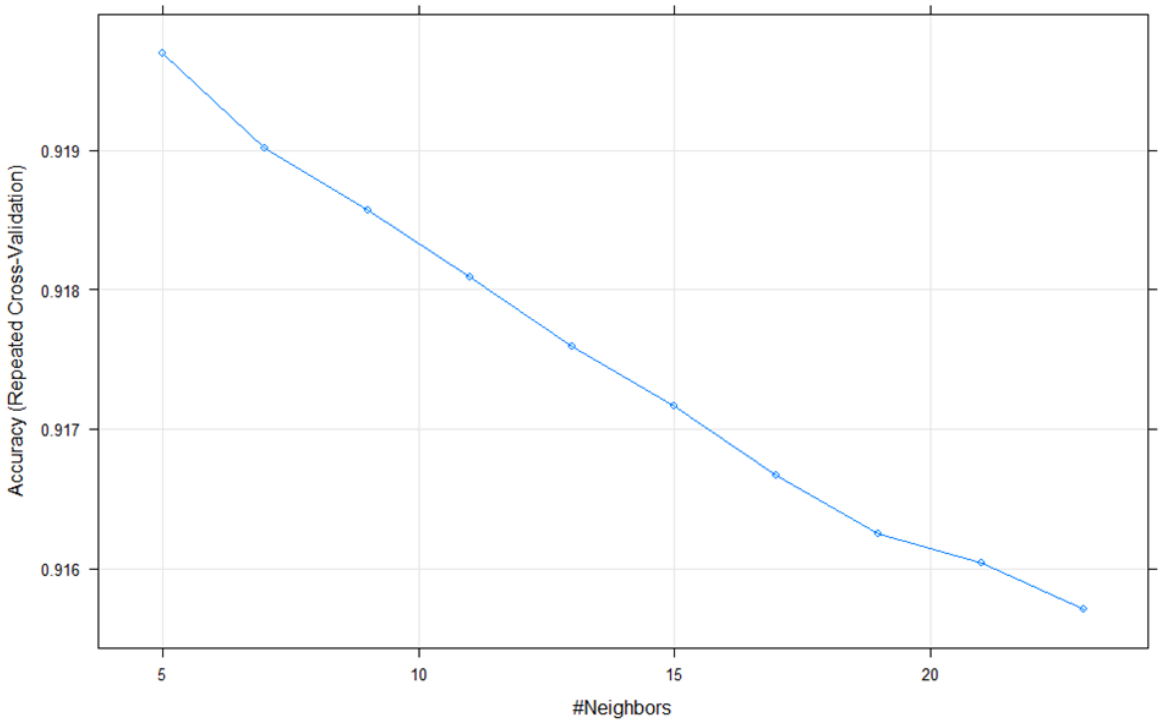
- In exploratory analysis, temperature proved to be an interesting variable when it came to organizational success (right)
- Browser Type, Device Type, OS Family and Temperature were chosen as predictor variables for model training
- 70% of the data was used to train the model; the remaining 30% tested the model
- Performance was assessed using Sensitivity, Specificity, and associated 95% CIs
- Organizations with identities not shown (???) were excluded from analysis—though a good next step would be testing the model on these events.



Model Results

Model training maximized accuracy at k = 5

Meaning only the closest 5 identified points in multidimensional space are used to determine a new point's organizational identity. Model accuracy decreased consistently with increasing k (below)



Classification Accuracy Varied Some

- Sensitivity was very high for Organizations 2 and 3; giving the model good ability to predict their success at auction
- Specificity was high for Organization 1; helping them know when not to bid

Overall Classification Accuracy was 92%

- 95% Confidence Interval: 91.9-92.2

	Organization 1	Organization 2	Organization 3
Sensitivity	.4889	.9611	.9746
Specificity	.9717	.7553	.9925
Positive Predictive Value	.6275	.9430	.9374
Negative Predictive Value	.9513	.8216	.9971
Prevalence	.0887	.8082	.1031
Detection Rate	.0434	.7768	.1005

Main Takeaways

1. Organization 3 was incorrectly classified as 1 just 33 times (right, top left corner), demonstrating excellent organizational discrimination.

2. With 92% classification accuracy, this model gives us the ability to take a user's information and adjudge which advertiser should bid on an ad.

3. Speculatively, organization 3 is likely the college since it differs so greatly from organizations 1 and 2, which seem to be more similar.