


A Comparison of Various Metrics Across Two Brain Supplement Versions

Sophia Leung
Springboard 2018

A dark blue diagonal gradient bar that starts from the bottom left and extends towards the top right, covering the lower half of the slide.

What are Nootropics?

- Cognitive enhancers or Smart Drugs
- Improves cognitive function
 - Memory
 - Focus
 - Motivation
 - Learning ability
- Can range from Caffeine to Omega-3 Fatty acids to Adderall



Objective

In order to determine which trial version is better the Overall Experience Rating will be used as the success metric:

- Each metric will be cross-compared to the overall experience rating, success will be determined if there is a correlation or likelihood of a user by its given metric will give a higher overall experience rating.
- For each metric, if the proportion of users in one trial is higher, then the metric is considered better for that respected trial.
- Classify users by their overall experience rating based on submetrics

Who are the clients?

The dataset is from my mentor Bill Scuba.
The name of the company was anonymized.

But could be used for:

- Pharmaceutical Companies
- Nutritional Supplement Manufacturers

Data Wrangling

Trial 99:

- 102 observations
- 40 unique users

Trial 51:

- 133 observations
- 47 unique users

- Categorical Ordinal Data
- Missing values imputed with mode if user answered for 2 other data collection times
 - If user did not answer for other times then it was left empty

Statistical Analysis

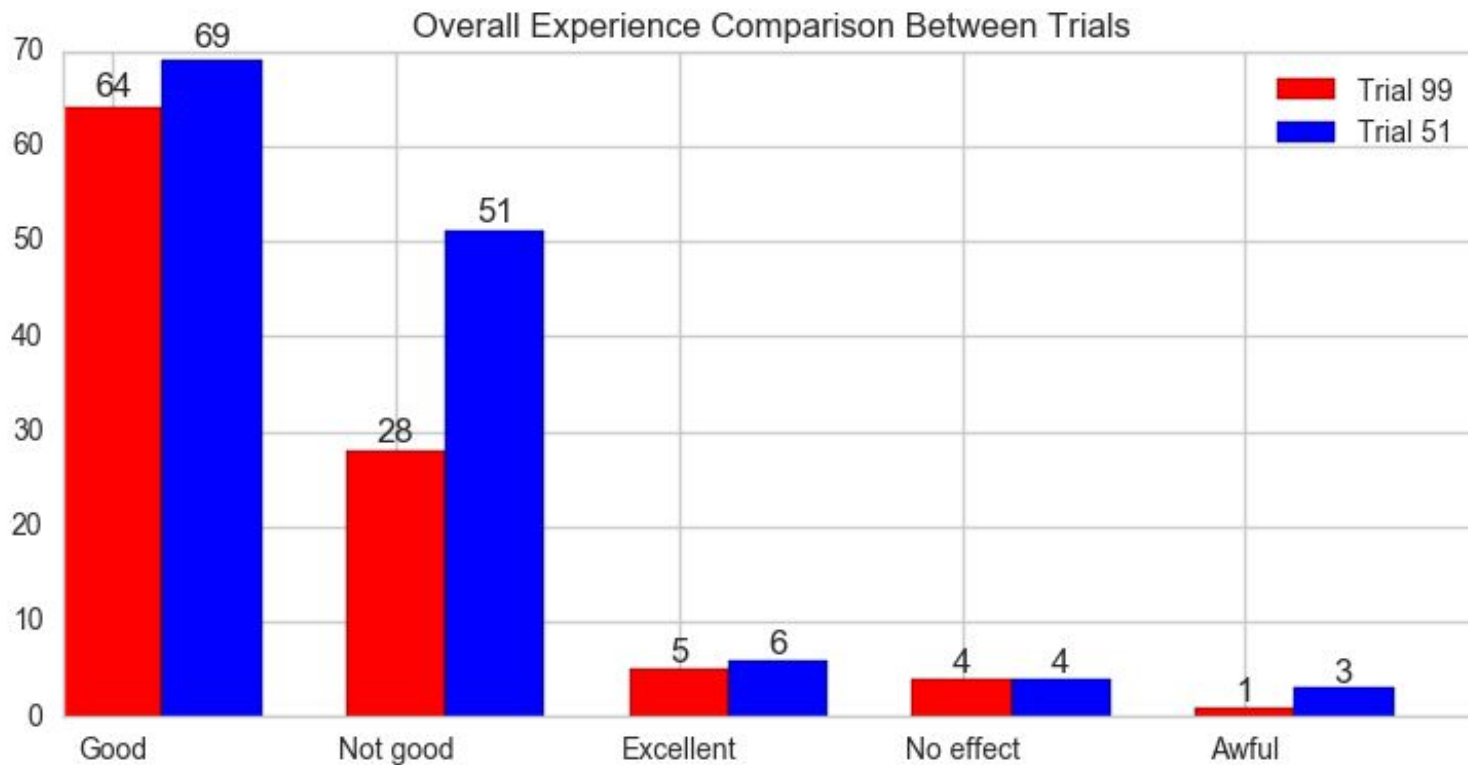
- Mann-Whitney U Test:
 - Compare if the median ranking between the two trials are different for given metric
- Overall Experience converted into “High” or “Low” rating
- Odds Ratio:
 - Compare between trials which metric are more likely to give a higher score
 - Compare within trials for that given metric for the likelihood of a higher overall score
- Linear-By-Linear Association:
 - If the metric for the given trial is correlated with the overall score

Statistical Analysis continued

- Two Proportions Test to compare between Trials
 - Grouped by ordinal rating
- Cohen's H to quantify effect size

Version	Version 51	Version 99
Sleep Quality		
Excellent	47	47
OK	72	49
Poor	13	4

Overall Experience



Overall Experience

Grouped ordinal ratings
into binary outcomes:

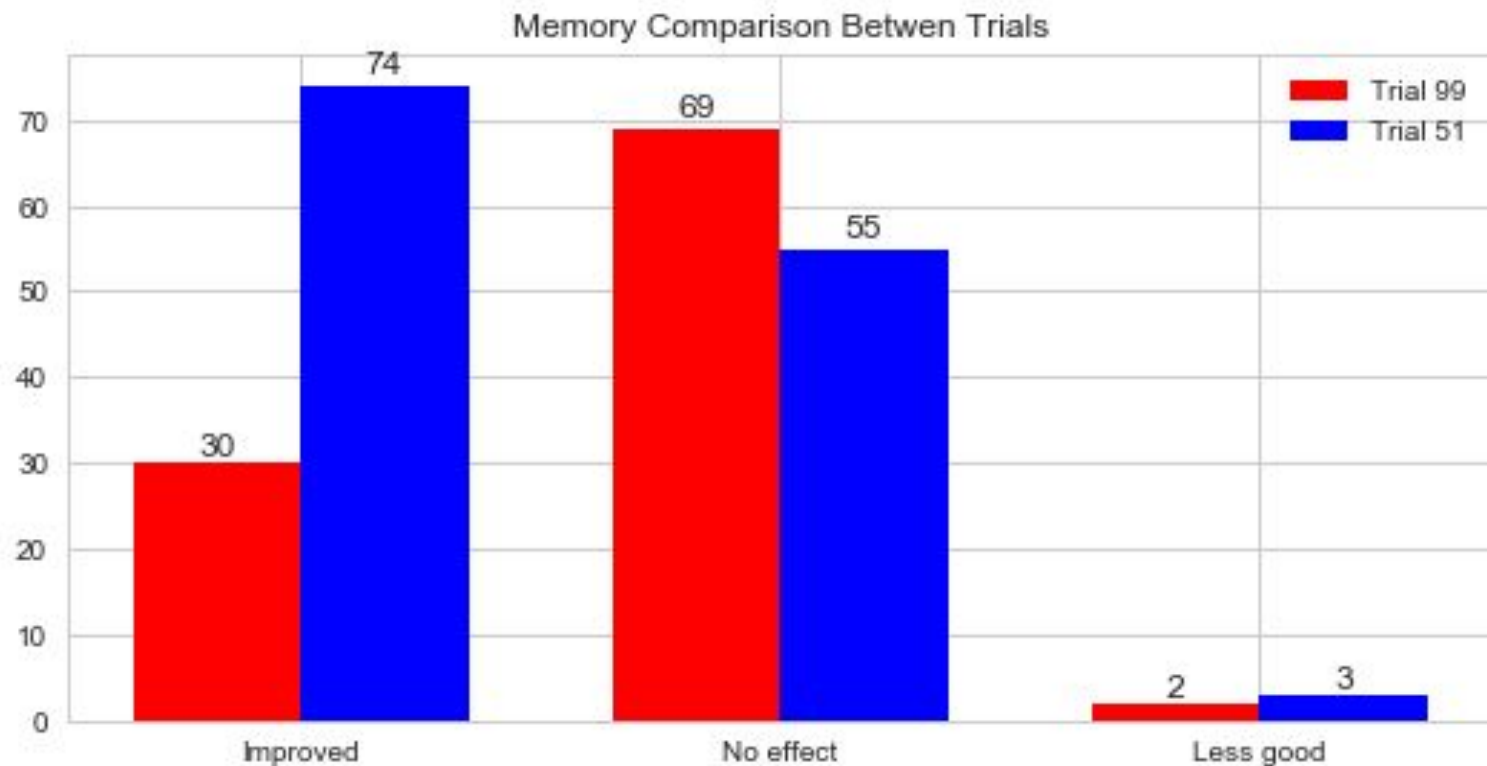
Version	Version 51	Version 99	All
High	120	92	212
Low	13	10	23
All	133	102	235

Excellent, good ➡ High

No effect, not good, awful ➡ Low

There's no difference between the proportion of users who gave a "high" rating between the two trials.

Memory

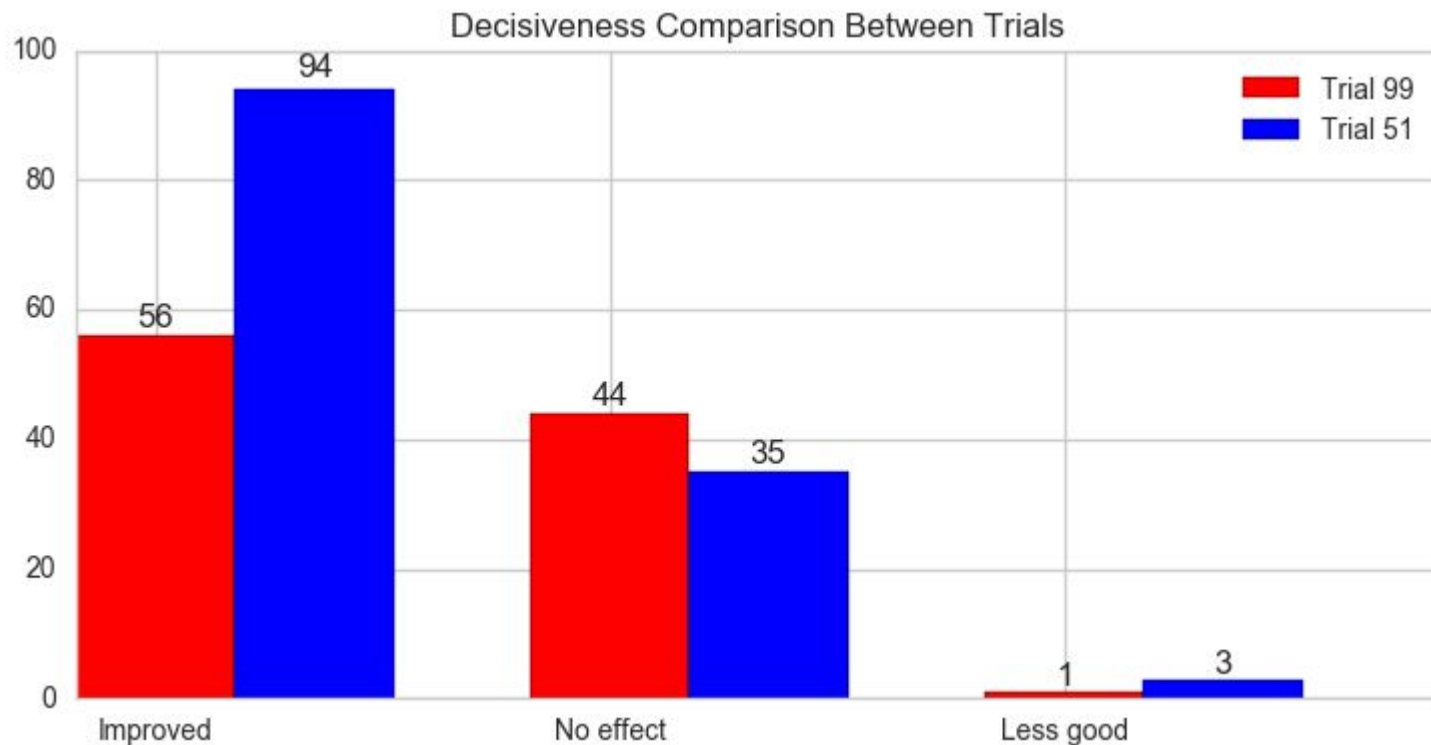


Memory

- Improved: 26.4% higher in Trial 51
- No Effect: 26.7% higher in Trial 99
- Correlation of improved memory and high overall experience in Trial 51
- Difference in median ranking

Version	Version 51	Version 99	All
Improved	74	30	104
Less good	3	2	5
No effect	55	69	124
All	132	101	233

Decisiveness

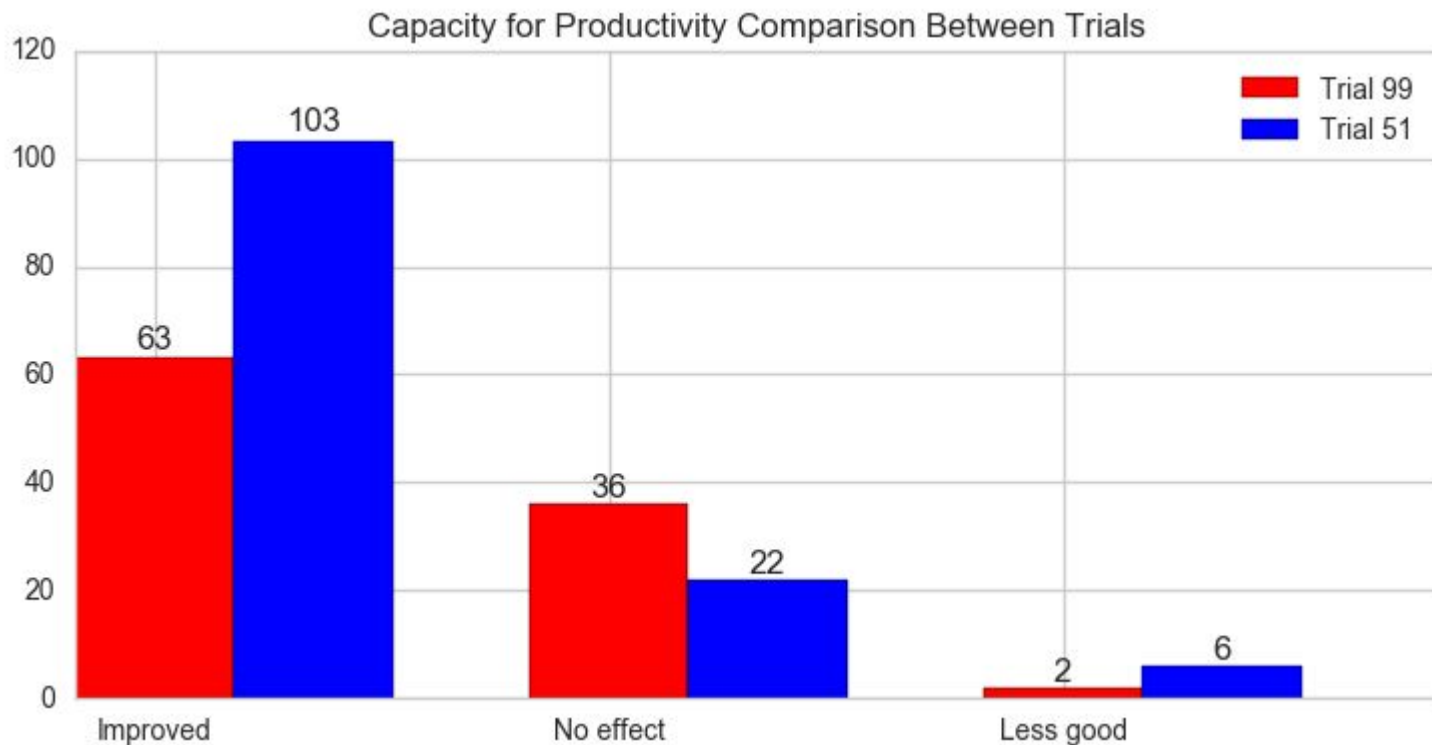


Decisiveness

- Difference in median ranking between trials
- Improved: 15.8% higher in Trial 51
- No Effect: 17% higher in Trial 99
- Both trials have a correlation between Decisiveness and a higher Overall Experience



Capacity for Productivity



Capacity for Productivity

- Difference in median ranking between the two trials
- Improved: 16.2% higher in Trial 51
- Both trials are correlated with Capacity for Productivity and a higher overall rating



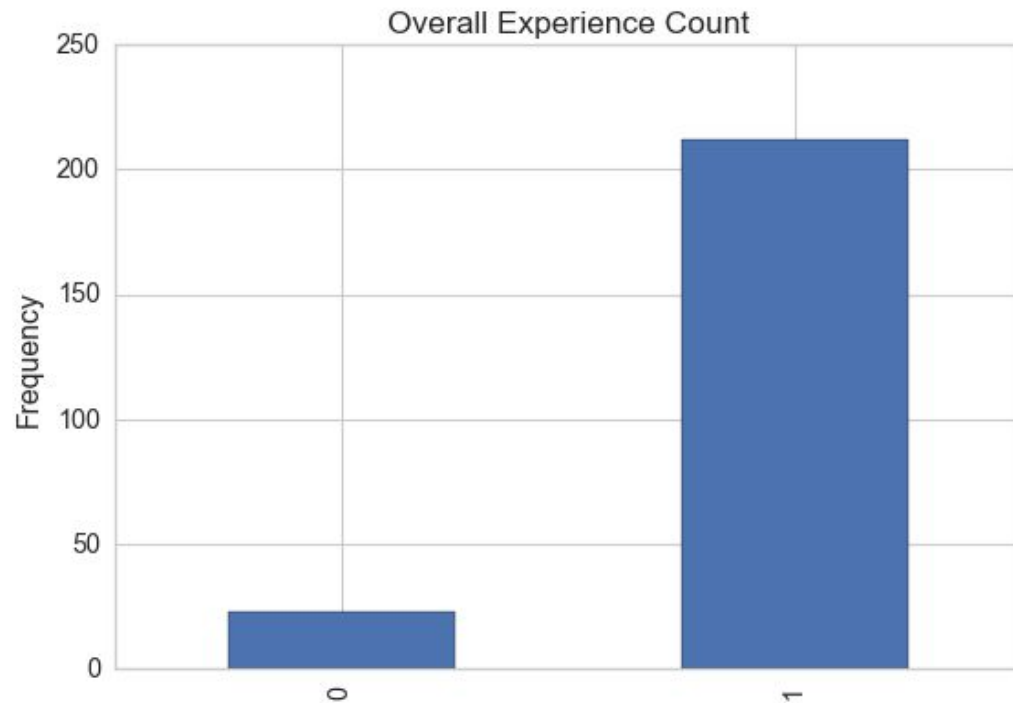
Machine Learning

Model Choice:

- SVM
- Logistic Regression
- Random Forest Classifier
 - Binary Classification
 - Multiclass
- Extra Trees Classifier

- Combined Both DataFrames together
- Converted categorical data into numeric with ordering

Machine Learning



Machine Learning: Feature Selection

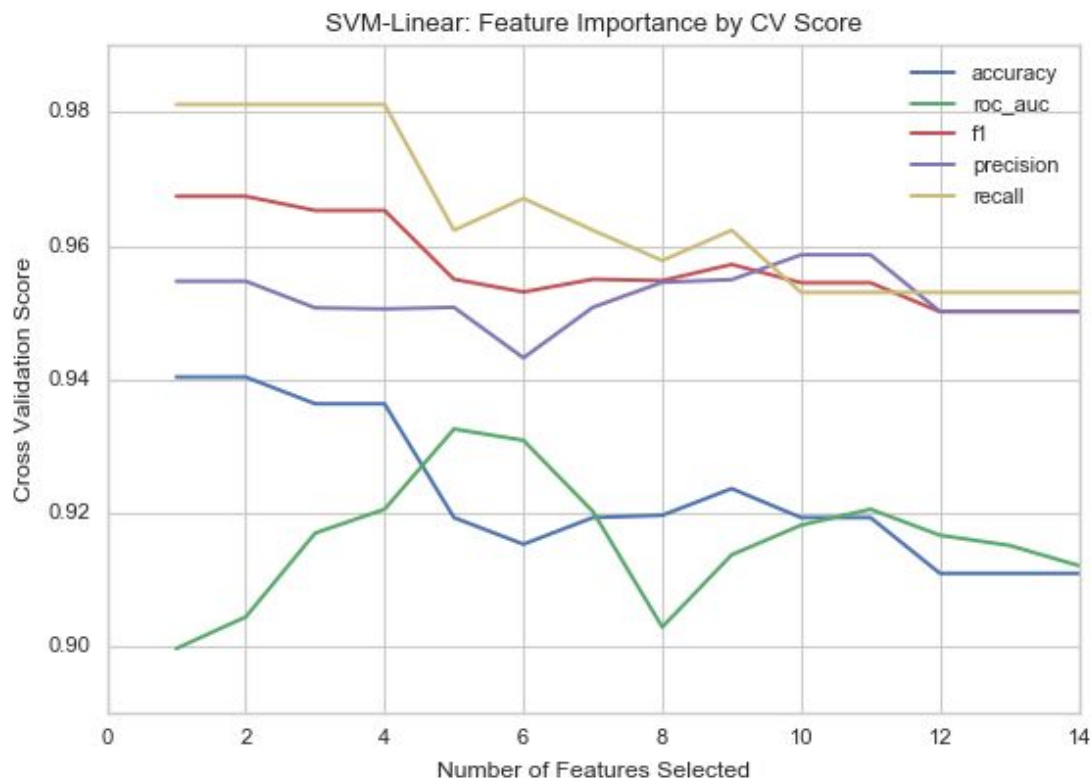
Wrapper Method: Recursive Feature Elimination with Cross-Validation for scoring metrics

- Accuracy, Precision, Recall, ROC Auc, F1

Optimal Number of Features

Model	Accuracy	ROC AOC	F1	Precision	Recall
SVM-Linear	2	5	2	11	4
Logit	14	8	14	14	14
Random Forest	8	12	1	13	2
Extra Trees	14	2	14	7	2

Machine Learning: Feature Selection



Machine Learning: Feature Selection

Recall has the highest score for all models. These are the important features based on recall

SVM	Logit	Random Forest	Extra Trees
Drive/Passion/Motivation	All 14 features	Physical Energy	Physical Energy
Focus/Atten/Concen			Focus/Atten/Concen
Physical Energy			
Interpersonal Cap			

Machine Learning: Class Imbalance

- Class imbalance issue was addressed using the SMOTE algorithm to up-sample the minority class
- Up-sampled only the training data

	Count	Up-Sampled
1	148	148
0	16	148

Machine Learning:

Recall	Training	Test
SVM	99.32%	95.77%
Logistic Regression	99.32%	88.73%
Random Forest	100.00%	90.14%
Accuracy	Training	Test
SVM	99.66%	94.37%
Logistic Regression	98.65%	84.51%
Random Forest	97.56%	94.37%
Random Forest (Multiclass)	63.41%	66.20%
Precision	Training	Test
SVM	100.00%	92.96%
Logistic Regression	100.00%	84.51%
Random Forest	97.52%	95.77%

Conclusion

Higher improved users in Trial 51
for:

- Memory
- Decisiveness
- Focus/Attention/Concentration
- Verbal Recall
- Capacity for Productivity

Higher improved users in Trial
for:

- Sleep Quality