Skill Craft Analysis

STAT 448

Final Report

Avinash Balakrishnan
Group 8

Introduction

The following analysis is based on StarCraft II game data. To begin with, StarCraft II is a Real-time Strategy game (RTS) video game with a 1 Vs. 1 setting. The overall aim of the game is to gather resources, develop a base and units, and using them in defeating the opponent by destroying their base. In the hands of professionals, the game is played in tremendous speeds. This combined with other factors such as the existence of an online leaderboard segregating players into different ranks based on their level of expertise, provides a unique pool of data that enables researchers to carry out various studies.

To further elucidate the nature of the game, StarCraft is played in a 'Top-Down" view. The games start with players choosing one of three species, each with its set of unique characteristics, further going on to build resources, units and building, establish a strategy and responding to opponents moves. One of the primary reasons for StarCraft's immense popularity is its emphasis on balanced gameplay by not giving any undue advantage to the players and ensuring all games are decided purely based on player skills.

Motivation of Study

The purpose of this particular dataset was to study human cognition; specifically on what separates experts from novices. One of the main factors that are studied is attention to our surrounding, which is an essential component to human learning process, and an important metric to this is human gaze, which is found to affect our learning. Given the diverse pool of players of the game and an established and tested ranking system, this provides a unique pool of data for researchers to use in these types of study. This provides an opportunity for better understanding of cognitive processing that take place in dynamic resources management scenarios and how people of different expertise handle such scenarios. From my analysis, I hope to effectively find a way to discriminate/differentiate between players of different levels of expertise based on the available data.

The Dataset

It is a collection of data from 3395 unique games played. The data was extracted from recorded game replays. Every observation in the dataset is identified with a unique and non-missing Game ID. The data consists of 20 variables including Game ID. These variables pertain to various for each game such as Basic Descriptive data, on screen movement, shortcut assignment and unit related. Of the 20 observations 15 were of continuous nature and were used as the variables used to do the analysis. The variable being predicted is League Index, which is ordinal with values 1-8.

An important factor in this dataset is a quantity called Perception action cycles (PACs) which is defined as "fixations with at least one action" on screen. To further get a good understanding of the variables, the descriptive statistics for a few of them are as shown below:

- Actions per minute (APM): This variable quantifies the number of actions per minute that a player performs. Intuitively, this would be higher in professional players who are capable of performing multiple tasks in a small amount of time.
- Action Latency: This gives the average time for the players of a particular group between consecutive PACs. Again, based on general trend, this quantity is considerably lower for professional players, who switch between actions very fast.

League Index	**APM**	**Action Latency**
	Mean	**Mean**
Bronze	59.54	95.40
Silver	74.78	81.27
Gold	89.97	73.70
Platinum	105.85	64.79
Diamond	131.52	56.09
Master	158.68	48.95
Grand Master	189.56	40.34
All	117.05	63.74

Fig: Means of APM and Action Latency

This difference between the variables should be able to separate the league index groups

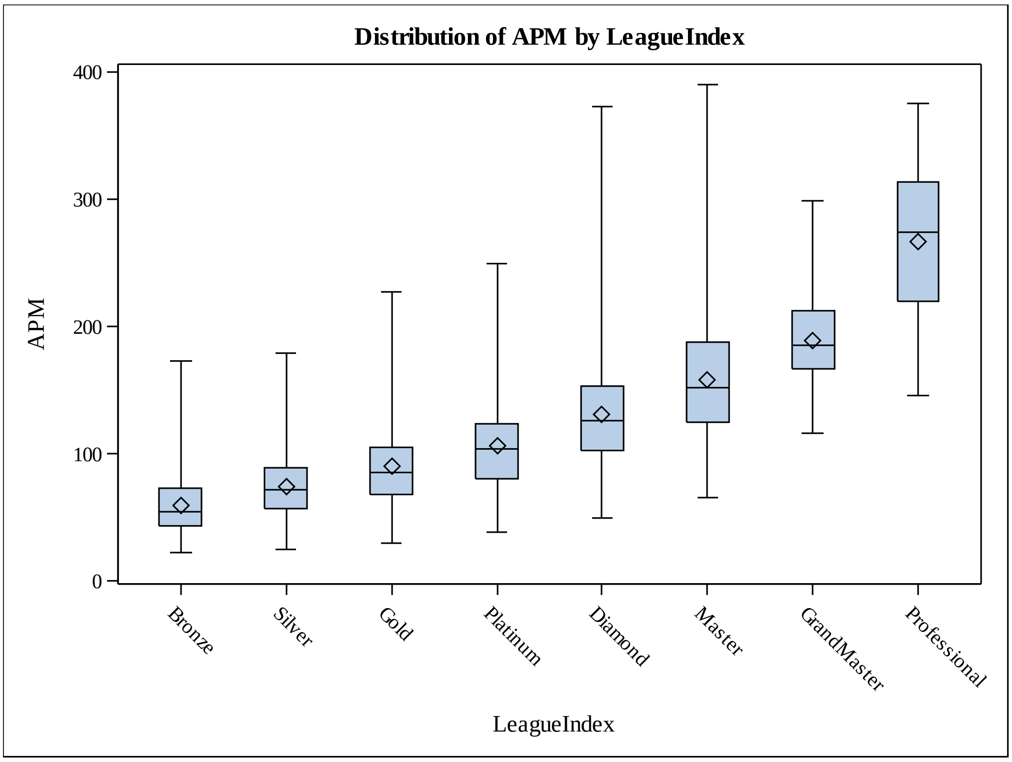

Fig: Box Plots for APM vs League Index

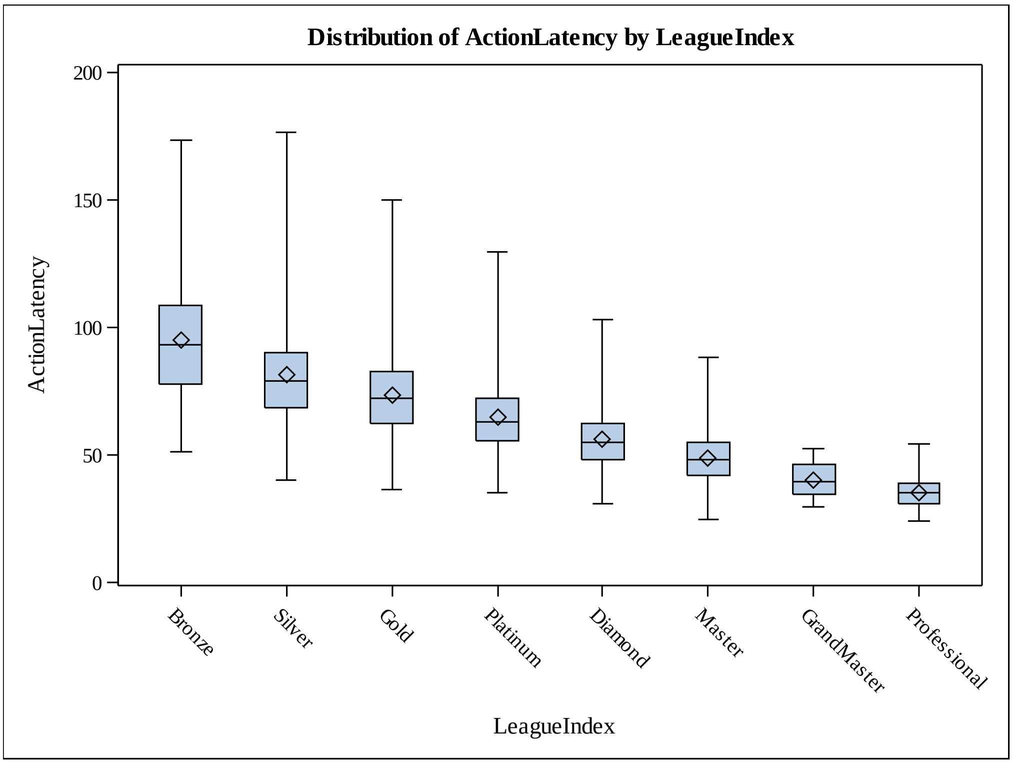

Fig: Box Plots for Action Latency vs League Index

Methods and Results

Data preparation:

All 3395 observations were used in the analysis with all 15 continuous variables used for predicting. The data was split into 2 sets, training and testing data of 70 % and 30 % of the total data. The data selection was done using simple random selection without replacement. The underlying proportions of the league index groups were maintained during the selection as shown in fig 1.

Master Dataset		
League Index	**Frequency**	**Percent**
Bronze	167	4.92
Silver	347	10.22
Gold	553	16.29
Platinum	811	23.89
Diamond	806	23.74
Master	621	18.29
Grand Master	35	1.03
Professional	55	1.62

Training Dataset		
League Index	**Frequency**	**Percent**
Bronze	117	4.92
Silver	243	10.21
Gold	388	16.30
Platinum	568	23.87
Diamond	565	23.74
Master	435	18.28
Grand Master	25	1.05
Professional	39	1.64

Fig 1: Frequency tables for Master and Training Datasets

Variable Selection

From the initial model based on all 15 predictors, a stepwise selection process was done to select the most significant of variables. The stepwise selection process had cutoff values of 0.15 for both entry and leaving, catering for the large dataset. The selection process was done based on the master dataset; in order to maintain the overall analysis with regard to the general population. This would ensure that variables that are highly insignificant would only be eliminated. Based on this, only 14 variables were found to be significant ('Complex abilities Used' was removed).

Model Training and Testing

Based on the variables selected above, discriminant analysis was run on the training model and the output was stored in a calibration dataset. This calibration data was used to test the test dataset and the error percent for both are as shown in the below table. The type of analysis used was Quadratic Discriminant analysis (QDA) with proportional priors.

Training Error count Estimate for League Index

	Bronze	Silver	Gold	Platinum	Diamond	Master	Grand Master	Professional	Total
Rate	0.273	0.794	0.757	0.448	0.773	0.485	0.320	0.102	0.602
Priors	0.049	0.102	0.163	0.238	0.237	0.182	0.010	0.016	

Testing Error count Estimate for League Index

	Bronze	Silver	Gold	Platinum	Diamond	Master	Grand Master	Professional	Total
Rate	0.440	0.875	0.806	0.493	0.800	0.559	0.900	0.562	0.671
Priors	0.049	0.102	0.163	0.238	0.237	0.182	0.010	0.016	

Fig: Training and Testing Error Counts Estimates Using QDA

The error percent observed for both the training and testing datasets are 60 and 67% respectively. This is a very high error rate. Though the error rate is reduced when using a constant prior proportion value, it does not make sense to have a constant prior proportions for a stratified dataset.

To eliminate the high error rates, the league indexes were classified into 3 groups. This was done based on the based on the high misclassification between adjacent groups. It was seen that the discriminant equation worked well to separate groups 1 and 8 but was not very good at discriminating between 4 and 5. This is shown in the error misclassification table in the appendices.

Number of Observations into League Index									
League Index	**Bronze**	**Silver**	**Gold**	**Platinum**	**Diamond**	**Master**	**Grandmaster**	**Professional**	Total
Bronze	**28**	**10**	**5**	6	0	1	0	0	50
Sliver	**48**	**13**	**17**	23	3	0	0	0	104
Gold	**32**	**22**	**32**	64	10	5	0	0	165
Platinum	25	12	26	**123**	**26**	31	0	0	243
Diamond	5	6	12	**107**	**48**	60	1	2	241
Master	1	4	4	49	37	**82**	**7**	**2**	186
Grandmaster	0	0	0	1	2	**6**	**1**	**0**	10
Professional	0	0	0	0	1	**8**	**0**	**7**	16
Total	139	67	96	373	127	193	9	11	1015
Priors	0.049	0.102	0.163	0.239	0.237	0.183	0.011	0.016	

Fig: Misclassified observations count for Test data (Model 1)

The above table gives the number of observations that were classified from/into each League Index. From the highlighted observations above, we can see that there is a trend i.e. observations from adjacent league indexes tend to get misclassified into one another. The same trend is also observed in the training data as shown in Appendices.

To eliminate the high error rates, the league indexes were classified into 3 groups. This was done based on the based on the high misclassification between adjacent groups. It was seen that the discriminant equation worked well to separate groups 1 and 8 but was not very good at discriminating between 4 and 5.

Model Refitting, Training and Testing

The model was refitted with the following groups as the response:

1. Bronze, Silver and Gold (Beginners)
2. Platinum and Diamond (Amateurs)
3. Master, Grand Master and Professional (Professionals).

This is done to ensure lower misclassification among adjacent groups. The model used was the same as the one in the previous section i.e. a Quadratic discriminant analysis with a proportional priors. The resulting error rates are substantially reduced and lesser observation were misclassified in this model. The response variable in this table is 'gpindex' which corresponds to the groups (Beginners, Amateurs, Professionals) that the league indexes were grouped into above.

Error Count Estimates for gpindex				
	1	**2**	**3**	**Total**
Rate	0.2273	0.4192	0.4669	0.3689
Priors	0.3143	0.4761	0.2097	

Fig: Error rates For Training Data with gpindex

Error Count Estimates for gpindex				
	1	**2**	**3**	**Total**
Rate	0.2759	0.4277	0.4953	0.3941
Priors	0.3143	0.4761	0.2097	

Fig: Error rates For Test Data with gpindex

The reduced error rates of 36.89 % for the training data and 39.41 % for testing data show certain improvement in the model.

Conclusions

From the above analysis, there is a reasonably good model that can classify the data. The initial model based on the League Indexes had a very high error rate with a training error rate of about 60.2 % and a testing error rate of about 67.1%, but the misclassification table of the model showed trends which indicated that adjacent groups are more likely to get misclassified between each other. This further motivated the use of a model based on grouped League indexes i.e. grouping adjacent league indexes into 3 groups, using a variable called 'gpindex'.

After doing further analysis using this grouping variable 'gpindex', it was found that the Error rate for the training was found to be 36.89 % and error rate for testing was found to be 39.41 %, which shows an improvement from the previous model. The initial proposal of classifying based on the League Index was not possible as the differences between the 8 league groups were not significant enough. It was however found that classification worked well when the observations were classified into 3 groups of Beginners, Amateurs and Professional.

Appendix

Appendix 1: Variable List and Definition:

1. Descriptive Data:
    - **Game ID:** Unique ID number for each game (Unique Integer)
    - **League Index**: Bronze, Silver, Gold, Platinum, Diamond, Master, Grand Master, and Professional leagues coded 1-8 (Ordinal)
    - **Age**: Age of each player (integer)
    - **Hours Per Week**: Reported hours spent playing per week (integer)
    - **Total Hours**: Reported total hours spent playing (integer)

2. On Screen Movement:
    - **APM**: Action per minute (continuous)
    - **Mini-map Attacks**: Number of attack actions on Mini-map per timestamp (continuous)
    - **Mini-map Right Clicks**: number of right-clicks on Mini-map per timestamp (continuous)
    - **Number Of PACs**: Number of PACs per timestamp (continuous)
    - **Gap Between PACs**: Mean duration in milliseconds between PACs (continuous)
    - **Action Latency**: Mean latency from the onset of a PACs to their first action in milliseconds (continuous)
    - **Actions In PAC**: Mean number of actions within each PAC (continuous)
    - **Total Map Explored**: The number of 24x24 game coordinate grids viewed by the player per timestamp (continuous)

3. Hotkey Assignment:
    - **Select By Hotkeys**: Number of unit or building selections made using hotkeys per timestamp (continuous)
    - **Assign To Hotkeys**: Number of units or buildings assigned to hotkeys per timestamp (continuous)
    - **Unique Hotkeys**: Number of unique hotkeys used per timestamp (continuous)
4. Unit Related:
    - **Workers Made**: Number of SCVs, drones, and probes trained per timestamp (continuous)
    - **Unique Units Made**: Unique unites made per timestamp (continuous)
    - **Complex Units Made**: Number of complex units trained per timestamp (continuous)
    - **Complex Abilities Used**: Abilities with specific targeting instructions used per timestamp (continuous)

Perception Action Cycles (PACs): are defined as "fixations with at least one action" on screen.

Hotkeys: Keyboard shortcuts assigned and used by players for more precise and faster game control.

Appendix 2:

Model Analysis and testing:

Form the initial **full model**, the MANOVA table tells us that at least one of the variables significantly predicts the group (p values <0.05)

Multivariate Statistics and F Approximations					
S=7 M=3.5 N=1685.5					
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
Wilks' Lambda	0.31788	40.33	105	21599	<.0001
Pillai's Trace	0.82444	30.07	105	23653	<.0001
Hotelling-Lawley Trace	1.71640	55.11	105	15660	<.0001
Roy's Greatest Root	1.43926	324.22	15	3379	<.0001
NOTE: F Statistic for Roy's Greatest Root is an upper bound.					

Fig: MANOVA table for full model

Based on the above table, a stepwise selection process was done to find the optimal subset of variables that could explain the data significantly.

Model Selection and Fitting

Model selection was done in a stepwise manner using the StepDisc statement with entry and leaving cutoff's at 0.15, catering for the large dataset. The below table gives the selection summary of the variables.

Stepwise Selection Summary										
Step	**Number In**	**Entered**	**Removed**	**Partial R-Square**	**F Value**	**Pr > F**	**Wilks' Lambda**	**Pr < Lambda**	**Average Squared Canonical Correlation**	**Pr > ASCC**
1	1	**APM**		0.4751	437.97	<.0001	0.52489133	<.0001	0.06787267	<.0001
2	2	**ActionLatency**		0.1904	113.79	<.0001	0.42492665	<.0001	0.08997481	<.0001
3	3	**AssignToHotKeys**		0.0703	36.55	<.0001	0.39506521	<.0001	0.09569207	<.0001
4	4	**MinimapAttacks**		0.0523	26.69	<.0001	0.37439214	<.0001	0.09984294	<.0001
5	5	**ActionsInPAC**		0.0473	23.97	<.0001	0.35669992	<.0001	0.10522442	<.0001
6	6	**WorkersMade**		0.0227	11.22	<.0001	0.34860135	<.0001	0.10789462	<.0001
7	7	**NumberOfPACs**		0.0213	10.52	<.0001	0.34117086	<.0001	0.11072394	<.0001
8	8	**GapBetweenPACs**		0.0202	9.97	<.0001	0.33427102	<.0001	0.11255598	<.0001
9	9	**TotalMApExplored**		0.0118	5.77	<.0001	0.33032479	<.0001	0.11370708	<.0001
10	10	**SelectByHotkeys**		0.0096	4.68	<.0001	0.32715436	<.0001	0.11456037	<.0001
11	11	**UniqueHotkeys**		0.0083	4.05	0.0002	0.32442919	<.0001	0.11536840	<.0001
12	12	**MinimapRightClick**		0.0069	3.36	0.0014	0.32218214	<.0001	0.11623928	<.0001
13	13	**ComplexUnitsMade**		0.0066	3.21	0.0022	0.32005288	<.0001	0.11705447	<.0001
14	14	**UniqueUnitMade**		0.0047	2.26	0.0271	0.31855915	<.0001	0.11750644	<.0001

Fig: misclassification observations for Training data

The data was fitted based on the above selected variables on the training dataset (**Model 1**). The Test for homogeneity of within covariance matrices is shown below.

Chi-Square	DF	Pr > ChiSq
8207.319360	840	<.0001

Fig: Test of Homogeneity of Within Covariance Matrices

Since the P value is <0.1, the assumption of pooled variance does not hold and we choose to use a Quadratic Discriminant analysis

Number of Observations into League Index									
League Index	**Bronze**	**Silver**	**Gold**	**Platinum**	**Diamond**	**Master**	**Grandmaster**	**Professional**	Total
Bronze	**77**	**19**	**8**	10	1	2	0	0	117
Sliver	**87**	**53**	**44**	48	11	0	0	0	243
Gold	**69**	**45**	**123**	116	25	10	0	0	388
Platinum	48	41	86	**295**	**56**	41	1	0	568
Diamond	17	19	44	**211**	**131**	136	5	2	565
Master	2	6	14	101	58	**229**	**15**	**10**	435
Grandmaster	0	0	0	1	1	**4**	**17**	**2**	25
Professional	0	0	0	0	0	**2**	**3**	**34**	39
Total	300	183	319	782	283	424	41	48	2380
Priors	0.049	0.102	0.163	0.239	0.237	0.183	0.011	0.016	

Fig: Misclassified observations count for Training data (Model 1)

The above table explains the reason for the high error rates observed. The highlighted numbers show the high misclassification trend between adjacent League Index groups. To address this issue the League Index was categorized into the following groups as a new variable **'gpindex'**.

1. Bronze, Silver and Gold (Beginners)
2. Platinum and Diamond (Amateurs)
3. Master, Grand Master and Professional (Professionals).

Based on the above grouping, the data was refitted with gpindex as the response (**Model 2**). The below table gives the class level information if gpindex:

Class Level Information					
gpindex	**Variable Name**	**Frequency**	**Weight**	**Proportion**	**Prior Probability**
1	1	748	748.0000	0.314286	0.314286
2	2	1133	1133	0.476050	0.476050
3	3	499	499.0000	0.209664	0.209664

Fig: Class level information gpindex (Model 2)

Number of Observations and Percent Classified into gpindex (Training)				
From gpindex	1	2	3	Total
1	578 77.27	154 20.59	16 2.14	748 100.00
2	335 29.57	658 58.08	140 12.36	1133 100.00
3	36 7.21	197 39.48	266 53.31	499 100.00
Total	949 39.87	1009 42.39	422 17.73	2380 100.00
Priors	0.31429	0.47605	0.20966	

Number of Observations and Percent Classified into gpindex (Testing)				
From gpindex	1	2	3	Total
1	231 72.41	86 26.96	2 0.63	319 100.00
2	151 31.20	277 57.23	56 11.57	484 100.00
3	10 4.72	95 44.81	107 50.47	212 100.00
Total	392 38.62	458 45.12	165 16.26	1015 100.00
Priors	0.31429	0.47605	0.20966	

Fig: Misclassified observations count for Training and Testing data (Model 2)

Error Count Estimates (Training)				
	1	2	3	Total
Rate	0.2273	0.4192	0.4669	0.3689
Priors	0.3143	0.4761	0.2097	

Error Count Estimates (Testing)				
	1	2	3	Total
Rate	0.2759	0.4277	0.4953	0.3941
Priors	0.3143	0.4761	0.2097	

Fig: Error rates For Training and Test Data with gpindex

The above table gives the outputs form the improved final model.

Acknowledgement

The original dataset if from the UCI Machine learning repository, and given below are the link to the dataset and the authors:

Source:

- Creators: Mark Blair, Joe Thompson, Andrew Henrey, Bill Chen
- Mark Blair: Department of Psychology; Simon Fraser University; Burnaby; 8888 University Drive; mblair '@' sfu.ca

Link:

https://archive.ics.uci.edu/ml/datasets/SkillCraft1+Master+Table+Dataset