Developing a Regression Algorithm for Predicting Magic: The Gathering Card Efficiency in Draft Format

Jonathan Tuck – jonathantuck@stanford.edu
Department of Electrical Engineering, Stanford University

Motivation

- **Magic: the Gathering** (Magic) is a **trading card game** that has controlled the genre with **over 20,000,000 players**.
- **Draft Format**: Open a pack, pick a card, and pass remaining cards. Repeat this process for three packs.
- In Draft, one of the most important factors in building a winning deck is **drafting "efficient" cards**, or cards that are relatively cheap (in terms of its mana cost) for the cumulative effect of the card on the game.

$$\text{Card Efficiency} = \frac{\text{Overall Positive Effect of Card}}{\text{Mana Cost}}$$

Problem Formulation

- Develop **regression algorithm** for determining card efficiency
  - Determine which base algorithm performs best
  - Alter base algorithm to fit Magic card rating framework
  - Incorporate unique Magic features
- Use **Draftsim** ratings as true values of card efficiency
  - Ratings used by professional Magic players in order to help card prioritization decisions
  - Ratings range from 0.0 to 5.0, in 0.1 increments
  - Inherent subjectivity in Draftsim ratings, but there **does exist** purely objective comparisons between cards

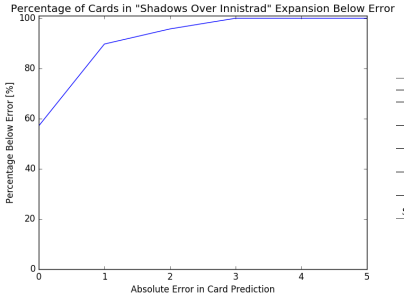
- Only features to be used are features derived from the Magic cards themselves

Features of a Magic Card

- **Mana Cost**
  - White, Blue, Black, Red, Green, or Colorless
  - Can have secondary mana costs
- **Card Type and Subtype**
  - Creatures, instants, sorceries, enchantments, artifacts, plainswalkers, and lands.
  - Important for "Tribal" effects – *e.g.,* "Destroy all non-Eldrazi Permanents."
- **Unique On-Card Effects**
  - Most complex feature, power typically proportional to number of words used to describe it.
  - Some mechanics are featured on multiple cards (left); others are featured on only one card (above)
- **Power/Toughness**
  - Metric of how fast it can damage opponent's board / keep itself from dying
- **Expansion Symbol**
  - Important in grouping cards of the same format
- **Card Name, Collector Information, Flavor Text, Artwork, Illustrator Name**
  - Purely aesthetic

Experimental Results

- Best results came from using **Multivariate Adaptive Regression Splines (MARS) with thresholding** to values between 0.0 and 5.0 in 0.1 increments.



	Percentage of Cards with Errors Under: [%]			
Error	**1.0**	**2.0**	**3.0**	**4.0**
Ordinary Least Squares	10	23	52	63
Locally Weighted Linear Regression	34	56	74	82
Logistic Regression	50	65	86	92
Multivariate Adaptive Regression Splines (MARS)	88	95	100	100
Locally Estimated Scatterplot Smoothing (LOESS)	16	23	63	73

- MARS makes predictions on ratings by fitting to the following form:

$$\hat{y}(x^{(i)}) = \sum_{i=1}^{m} a_i B_i(x^{(i)})$$

- $x^{(i)}$ is the numerical mapping of the card features, a_i are constants, and B_i are basis functions that are in general taken to be a combination of either hinge functions or the product of two or more hinge functions. Recall that a hinge function h(x) = max{0,x}.
- MARS with thresholding **classified ~60% of cards perfectly**, and 88% to within 1.0 rating units.
- Some lack of accuracy is to be expected: there exists subjectivity in the rating scheme.

Future Work

Predict Based on past draft choices:
- This would lead to predicting the most efficient, synergized deck.

Many more formats to explore:
- "Sealed" – Construct decks from 6 packs opened immediately prior to gameplay
- "Standard" – Construct decks from most recent expansions
- "Modern" – Construct decks from expansions from the last 10 years
- "Vintage" – Construct decks using any card ever made

References

[1] Wizards of the Coast. *Emrakul, the Promised End*. 2016. [Online]. Available: http://gatherer.wizards.com/Pages/Card/Details.aspx?multiverseid=414295. [Accessed: 10 October 2016].
[2] MTG JSON. 2013. [Online]. Available: https://mtgjson.com/. [Accessed: 10 October 2016].
[3] DraftSim. 2016. [Online]. Available: http://draftsim.com/. [Accessed: 1 November 2016].
[4] TensorFlow. 2016. [Online]. Available: https://www.tensorflow.org/. [Accessed: 3 November 2016].
[5] scikit-learn. 2016. [Online]. Available: http://scikit-learn.org/stable/. [Accessed: 9 November 2016].
[6] Ng, Andrew. *CS 229: Supervised Learning*. 2016. [Online]. Available: http://cs229.stanford.edu/notes/cs229-notes1.pdf. [Accessed: 10 November 2016].
[7] Friedman, Jerome H. *Multivariate Adaptive Regression Splines*. Ann. Statist. 19 (1991), no. 1, 1–67. doi:10.1214/aos/1176347963. http://projecteuclid.org/euclid.aos/1176347963.
[8] Cleveland, William S. (1979). *Robust Locally Weighted Regression and Smoothing Scatterplots*. Journal of the American Statistical Association. 74 (368): 829–836. doi:10.2307/2286407
[9] Ng, Andrew. *CS 229: Regularization and Model Selection*. 2016. [Online]. Available: http://cs229.stanford.edu/notes/cs229-notes5.pdf. [Accessed: 10 November 2016].
[10] Rudy, Jason. *A Python implementation of Jerome Friedman's Multivariate Adaptive Regression Splines*. 2016. [Online]. Available: https://github.com/scikit-learn-contrib/py-earth. [Accessed: 15 November 2016].