

Predicting the strength of Magic: The Gathering cards from card mechanics

Dustin Fink, Benjamin Pastel, Neil Sapra



Objectives

Develop a discriminative regression model to predict the market price of Magic: The Gathering cards.

Assess the value of a card by analyzing features intrinsic to the card such as its power, toughness, mana cost, and rules text.

Data

We obtained data for **15,708 unique cards** through the MTG JSON database [1]. Price data was obtained through MTGStocks [2].

Features were extracted through our **in-house feature extractor**.

Pathological card sets, characterized by their collectability, scarcity, and farcity, were removed.



```
{
  "colorIdentity": [
    "U"
  ],
  "imageName": "vendilion clique",
  "toughness": "1",
  "power": "3",
  "text": "Flash, Flying
  When Vendilion Clique enters the battlefield, look at target player's hand. You may choose a nonland card from it. If you do, that player reveals the chosen card, puts it on bottom of his or her library, then draws a card.",
  "subtypes": [
    "Faerie",
    "Wizard"
  ],
  "layout": "normal",
  "name": "Vendilion Clique",
  "manaCost": "{1}{U}{U}",
  "cmc": 3,
  "colors": [
    "Blue"
  ],
  "type": "Legendary Creature - Faerie Wizard",
  "supertypes": [
    "Legendary"
  ],
  "types": [
    "Creature"
  ]
}
```

Methods

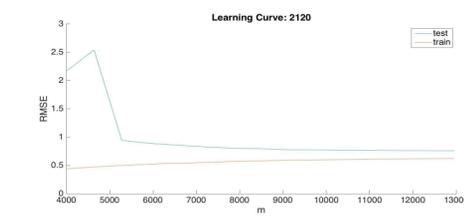
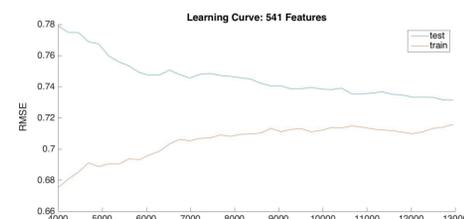
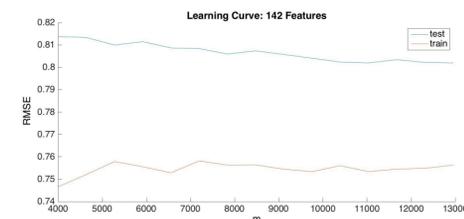
Linear regression on log-dollars

3-grams: Surprisingly effective!

- We strung together all the card text, and created features for each consecutive 1, 2, or 3 word phrase.

Regularization

- Drop features that are too rare.
- Drop features that are too similar.
- Determine threshold to optimize bias-variance trade-off



Other (less successful) methods:

- Locally weighted linear regression
- Crossing features
- Mutual Information regularization

Results

Metrics

- Training size: 12916 cards
- Testing size: 1436 cards
- Train RMSE: 0.712 log-dollars
- Test RMSE: 0.709 log-dollars
- IOR on test: 35.51%
- IOR on test: 35.76%

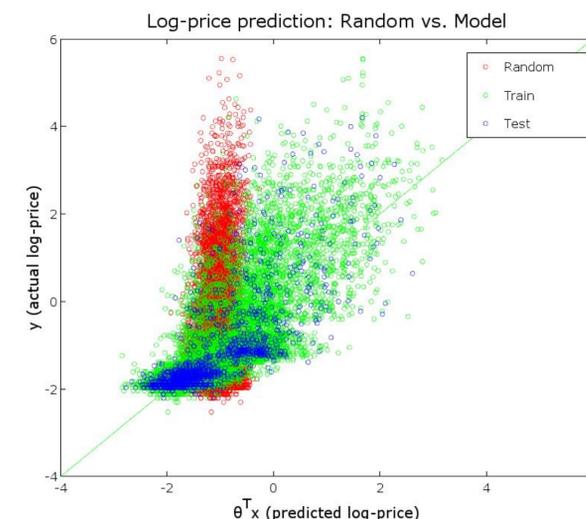
(IOR = Improvement over randomly generated features)

Significant features

Theta	Feature
1.048	Rare
1.014	"Additional"
0.7753	Mythic
0.5005	"search"
...	
0.2325	"Sacrifice a creature"
...	
-0.4863	Common
-1.035	"An additional"

Holding out new sets performance

Origins and Zendikar RMSE: 0.8



Conclusions

Feasibility

Despite the sparsity of the feature matrix and surprising ineffectiveness of crossing features like mana cost with rules, our model was able to find the variation in log-price. The model provides non-trivial classification ability on the \$1 boundary.

Gameplay Mechanics

The model was able to pick up on powerful in-game mechanics such as searching in the library, drawing cards, and destroying opponent's creatures.

Predictive Power

Training on all but the most recent 500 cards (Magic Origins and Battle for Zendikar) we achieve a RMSE of 0.8, indicating the usefulness of the model for predicting the value of newly printed cards.

References

- [1] Schults, et al. MTG JSON Database. <http://mtgjson.com/>
- [2] MTGStocks. <http://www.mtgstocks.com/>

Acknowledgements

We thank Sam Corbett-Davies for helpful advice and discussion. In addition, we are grateful to Professor Andrew Ng and the wonderful TA staff for the excellent course.