**Objective**

- Create a system that takes existing card data (past prices, tournament usage, card attributes) and predicts future prices
- Be able to determine whether a card is worth purchasing or selling in a given time frame

**Motivation**

- The market of the Magic the Gathering trading cards is a simpler and smaller version of the stock market
- Stock price prediction is a known and challenging problem: making progress on the simpler analogue could provide insight
- Players of the card game would like to know when it is best for them to buy and sell cards
Card Price Data for
Boris A. Perkhounkov, Cooper Gates Frye, E
Stanford University CS229: Machine

Data Extraction (cont.)

Tournament Deck Data
- "Pro Tours":
  - Are highly competitive tournaments
  - Take place after new sets come out (4 times a year)
  - Are a deciding factor in the prices of new cards
- Data from past 12 (overlaps with price data time period)
- Collected top 8 deck lists for Pro Tours, counted appearances

Want to know
- Use Support Vector Regression
  - Unsure of linearity
  - SVR is a comp...
Preprocessing
  - mean 0 and v...
One input for ahead we are
  - The price for the price, so the pr...
  - The price in two factors, and it
Each run split
  - 30% testing p...
We choose the evaluating wi...
Magic the Gathering

Marilyn Margaret Franklin

Learning

**Methodology**


Standardized features to have variance 1.

Our system is the number of days predicting tomorrow will usually be close to today’s prediction will appear to be too accurate. 10 years is affected by many random factors, is hard to make use of accurate predictions.

Split the data into 70% training and 30% test.

Tune hyperparameters by grid search, with 7-fold cross validation.

**Results**

- As expected, accuracy increases the greater the prediction horizon.
- The methods which rely on market trends and price outliers are more affected by the number of days ahead.
- Results in this plot are for 1500 predictions.
For all methods, decreases in time span rely more heavily on past ed.

e from a training set of size
Data Extraction

- 15440 distinct cards in the game
- Historical Price Data
  - Scraped from http://www.mtgstocks.com/
  - 26348 cards’ data (repetitions from reprints)
  - Price data ranges from June 9, 2012 - Nov. 6, 2015 (filled in with 0s)
  - Ignored cards that cost less than $1 (80%) - they aren’t worth it to buy/sell because of transaction costs, and they have very low variance in price, so they aren’t very predictive
Tournament legality is partially determined by set

- All prediction
- Tried $C = 1$,
- Linear kernel polynomial ke

Card Attributes Data
- Downloaded JSON files containing card information from mtgjson.com
- Parsed by in-game cost, keywords in description of usage
- Such attributes determine usefulness and play a large factor in demand

Accuracy decreases
- Cross validation helps alleviate this
30 days forward
10, 100, 500 with $C = 100$ the best performed better than radial or kernels

Varying Size of Training Set

- Most surprising result from the tournament
- We suspect there are at fault
  - The number of cards (es) tournament play is very
  - There are more tournament demand from what our o
- Decks beyond the top could help, but the da
- The other surprising accuracy was.
- May be a result of fin cards, where predicti
- Predicting those has l make a profit by buyi
- Next step is to redefi tracks cards with dyn
- Then we will use that further modify our fea
1 Further Work

was the lack of impact
data.
two main effects currently

p. expensive cards) that see
small.
ents played that gradually change
ata currently captures.

8 and smaller tournaments
ata is unavailable
result is how high our

ancially stable expensive
n is a trivial task
ow utility, since you can’t
ng them or selling them.
he accuracy so that it only
amic prices
new accuracy metric to
ture selection.