Predictir

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

ıg Card Price Data for

Boris A. Perkhounkov, Cooper Gates Frye, Ei

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


```
Tournament 1, Winner's main deck ..... 0
Tournament 1, Winner's side deck ..... 0
...
Tournament 5, Winner's side deck ..... 0
Tournament 5, other main decks ..... 5
Tournament 5, other side decks ..... 0
Tournament 5, all decks ..... 5
To   ment 6, Winner's main deck ..... 0
      6, Winner's side deck ..... 0
        her main decks ..... 7
      , other side decks ..... 3
To   ment 6, all decks ..... 10
Tournam ..... 1
Tournam ..... 0
Tournam ..... 10
Tournam ..... 0
Tournam ..... 11
Tournam ..... 0
...
Tournament 12, all decks ..... 0
```

Appearances in Decks

Set Membership Data

- Affects supply (older sets are not printed)

- Want to know
- Use Support V
  - Unsure of linea
  - SVR is a comp
- Preprocessing mean 0 and va
- One input for ahead we are
  - The price for t price, so the pr
  - The price in tw factors, and it
- Each run split 30% testing p
- We choose the evaluating wit

Magic the Gathering

mily Margaret Franklin

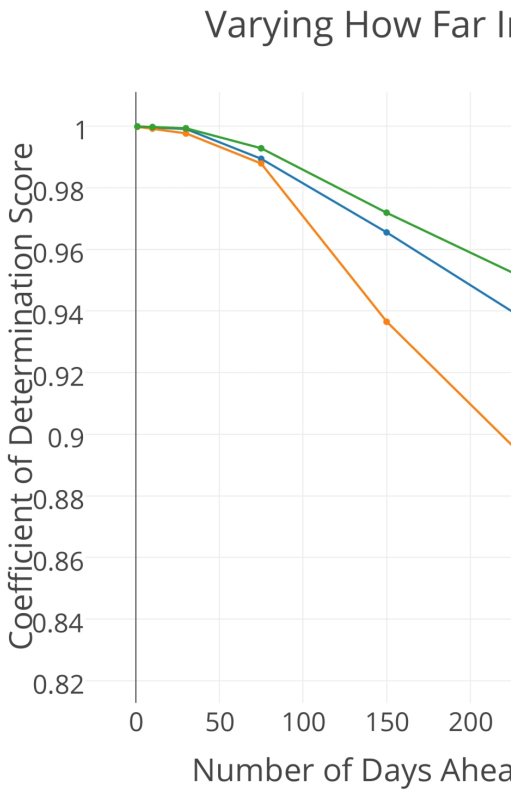
Learning

Methodology

future price: a regression problem.
Vector Regression
rity of the problem
utationally efficient method,
: standardized features to have
ariance 1
our system is the number of days
predicting
omorrow will usually be close to today's
ediction will appear to be too accurate.
/o years is affected by many random
is hard to make use of accurate predictions.
s the data into 70% training and
oints
e hyperparameters by grid search,
h 7-fold cross validation.

Results

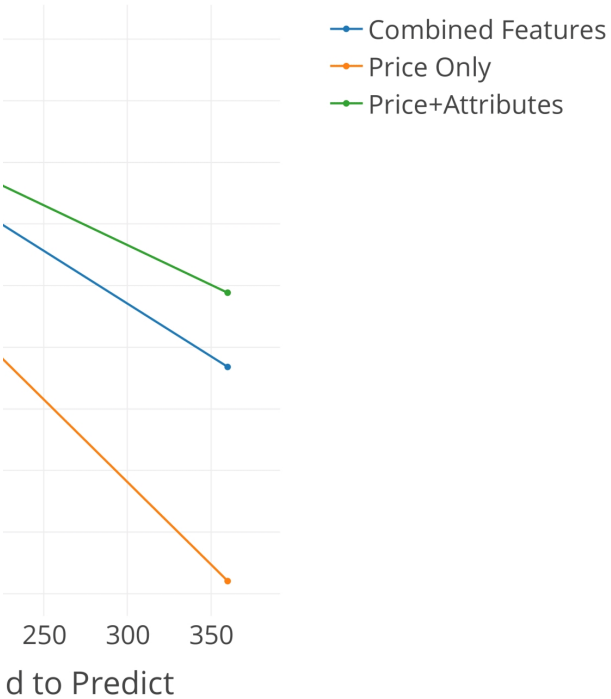
Resul



- As expected, accuracy
the greater the predic
- The methods which re
prices are more affecte
- Results in this plot ar

ts (cont.)

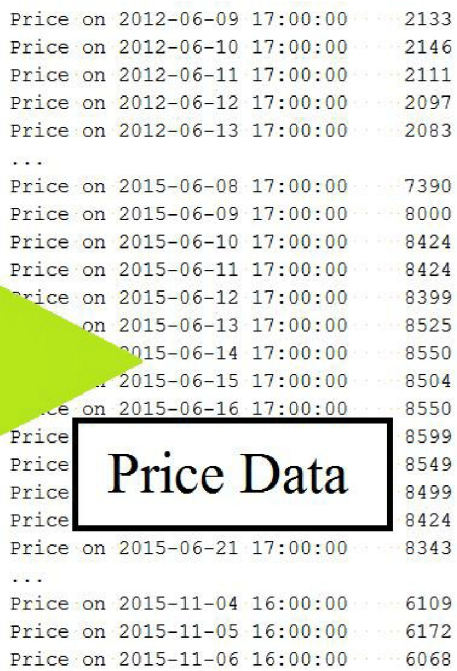
ıto the Future to Predict



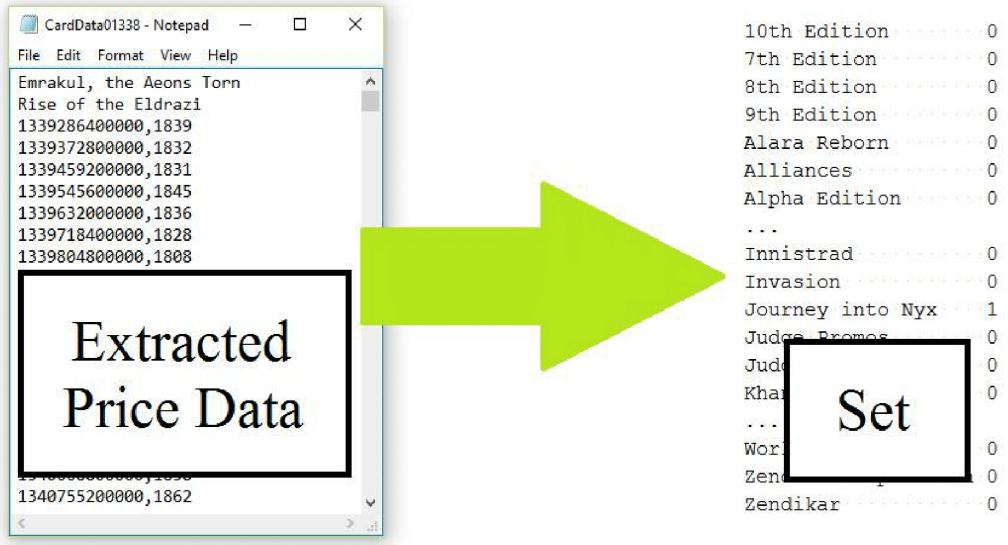
for all methods decreases
tion time span
ely more heavily on past
ed.
e from a training set of size

Data Extraction

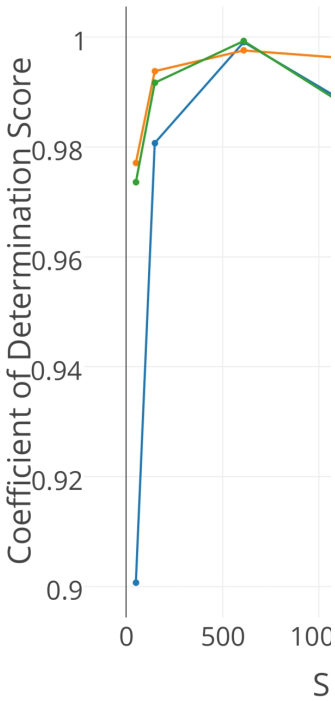
- 15440 distinct cards in the game
- Historical Price Data
  - Scraped from http://www.mtgstocks.com/
  - 26348 cards' data (repetitions from reprints)
  - Price data ranges from June 9, 2012 - Nov. 6, 2015 (filled in with 0s)
  - Ignored cards that cost less than $1 (80%) - they aren't worth it to buy/sell 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 deca
- Cross validati
  alleviate this

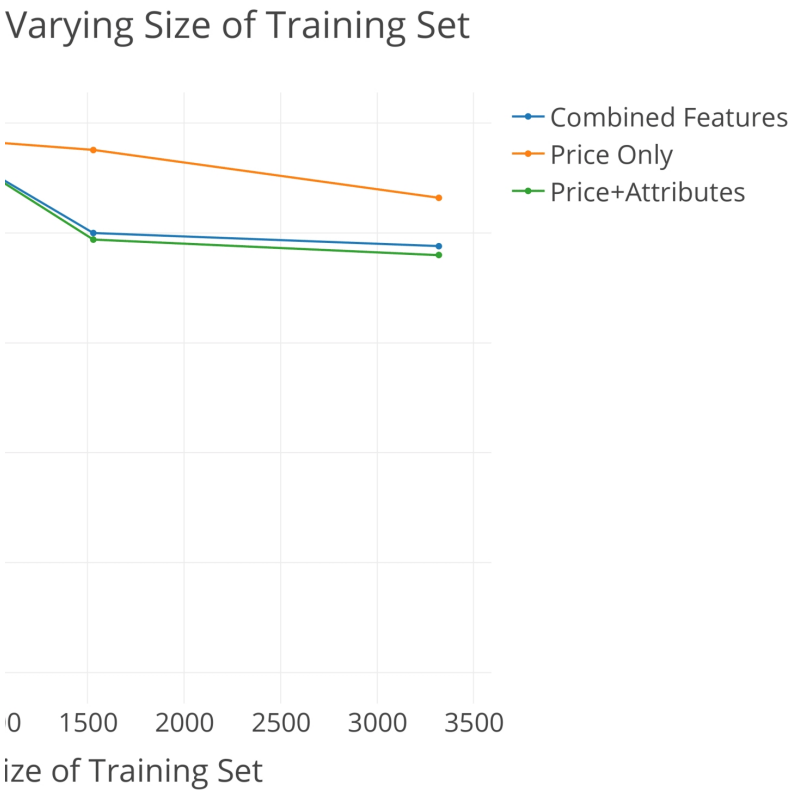
1500.

s 30 days forward

10, 100, 500 with $C = 100$ the best

performed better than radial or

rnels



ays unexpectedly

on with random sampling will help

artifact

Analysis and

- Most surprising result from the tournament
- We suspect there are at fault
  - The number of cards (es tournament play is very
  - There are more tournam demand from what our
- Decks beyond the top could help, but the da
- The other surprising r accuracy was.
- May be a result of fin cards, where predictio
- 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

l Further Work

was the lack of impact
data.
two main effects currently

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

8 and smaller tournaments
ıta is unavailable

esult is how high our

ancially stable expensive
n is a trivial task

ow utility, since you can't
ng them or selling them.

ıe accuracy so that it only
amic prices

new accuracy metric to
ıture selection.