# Corporate Default Prediction via Deep Learning

Shu-Hao Yeh,[1] Chuan-Ju Wang,[1] Ming-Feng Tsai[2]
[1]University of Taipei, Taipei, Taiwan
[2]National Chengchi University, Taipei, Taiwan

July 1, 2014

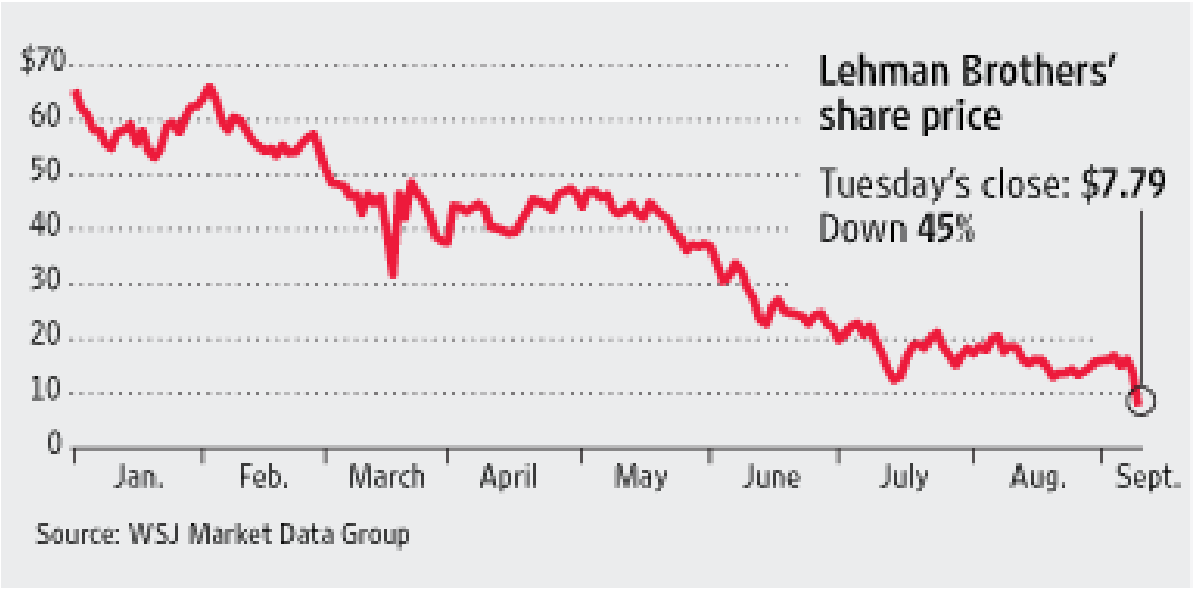
# Outline

# Outline

# Introduction

- Lehman Brothers

| Company | Filing date | Total Assets pre-filing | Assets adjusted to the year 2012 | Filing court district |
|---|---|---|---|---|
| Lehman Brothers Holdings Inc. | 2008-09-15 | $639,063,000,800 | $700 billion | NY-S |



# Introduction

1. Classical statistical models
   - Altman's model, Z-Score
   - Ohlson's model, O-Score
2. Market-based models
   - KMV-Merton model
3. Machine learning models
   - Support vector machines
   - Artificial neural network

# Introduction

- Feature Selection Problem
  - What are good features?
    1. 10-day moving average?
    2. Minimum stock price?
    3. Maximum stock price?
    4. Standard deviation of stock prices?

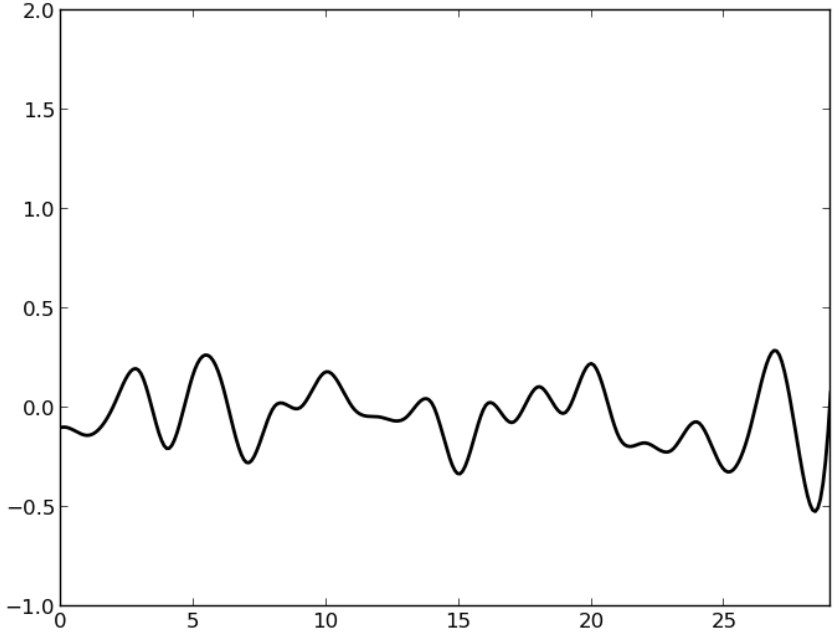
# Introduction

- Deep learning
  1. A machine learning method based on learning representations
     $\rightarrow$ Different concepts are learned from other concepts, with the more abstract, higher level concepts being learned from the lower level ones.
     $\rightarrow$ Deep learning helps to disentangle these abstractions and pick out which features are useful for learning.
  2. Architectures: deep neural networks, convolutional deep neural networks, and deep belief networks.
     $\rightarrow$ Applied on computer vision, automatic speech recognition and natural language processing

# Introduction

- Main contributions
  1. Transform discrete stock return time series to a graph representation

$$[-0.098684, -0.138686, 0.016949, \cdots, -0.365854, 0.076923]$$

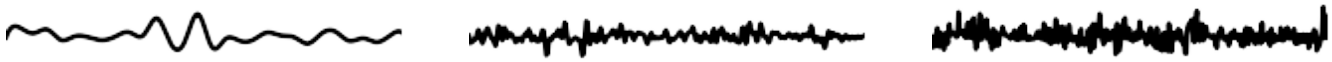


# Introduction

- Main contributions
  2. Adopt deep learning algorithms on the graphs for corporate default prediction
     - The stock returns of the default companies

(a) 30-days prior to default (b) 180-days prior to default (c) 360-days prior to default

     - The stock returns of the non-default companies



(d) 30-days (e) 180-days (f) 360-days

# Outline

# Stock Return Calculation

- Daily stock return

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}}$$

$$S_{t-1} = \text{stock price at day } t-1$$

$$S_t = \text{stock price at day } t$$

# Problem Description

- Default prediction can be treated as a classification problem.
  - Input: graphs of stock daily returns
  - Output: 0 (non-default) and 1 (default)
  - Algorithm: Deep Belief Network (DBN) (python toolkit: theano[1])

[1] http://deeplearning.net/software/theano/

# Outline

# Dataset

- The daily stock returns of American publicly-traded companies are from the Center for Research in Security Prices of Wharton Research Data Services.[2]
  - From year 2001 to 2011

| PERMNO | Date | Returns |
|---|---|---|
| 10001 | 20030102 | 0.088287 |
| 10001 | 20030103 | 0.052500 |
| 10001 | 20030106 | -0.019121 |
| 10001 | 20030107 | 0.004964 |
| 10001 | 20030108 | 0.006024 |

[2] https://wrds-web.wharton.upenn.edu/wrds/

# Dataset

| Year | # of all companies | # of default companies | Prior_30 | Prior_180 | Prior_360 |
|---|---|---|---|---|---|
| 2001 | 8608 | 982 | 982 | 964 | 398 |
| 2002 | 7900 | 706 | 704 | 694 | 671 |
| 2003 | 7475 | 606 | 606 | 600 | 588 |
| 2004 | 7475 | 449 | 449 | 446 | 437 |
| 2005 | 7364 | 489 | 486 | 480 | 469 |
| 2006 | 7423 | 468 | 468 | 460 | 441 |
| 2007 | 7679 | 602 | 601 | 595 | 581 |
| 2008 | 7394 | 553 | 551 | 542 | 502 |
| 2009 | 7141 | 517 | 514 | 509 | 489 |
| 2010 | 7085 | 450 | 449 | 442 | 425 |
| 2011 | 7112 | 404 | 403 | 395 | 381 |

1. Prior_30: # of default companies having 30-day daily stock returns prior to default
2. Prior_180: # of default companies having 180-day daily stock returns prior to default
3. Prior_360: # of default companies having 360-day daily stock returns prior to default

- For each year, we construct a balanced dataset for training.

# Experimental Settings

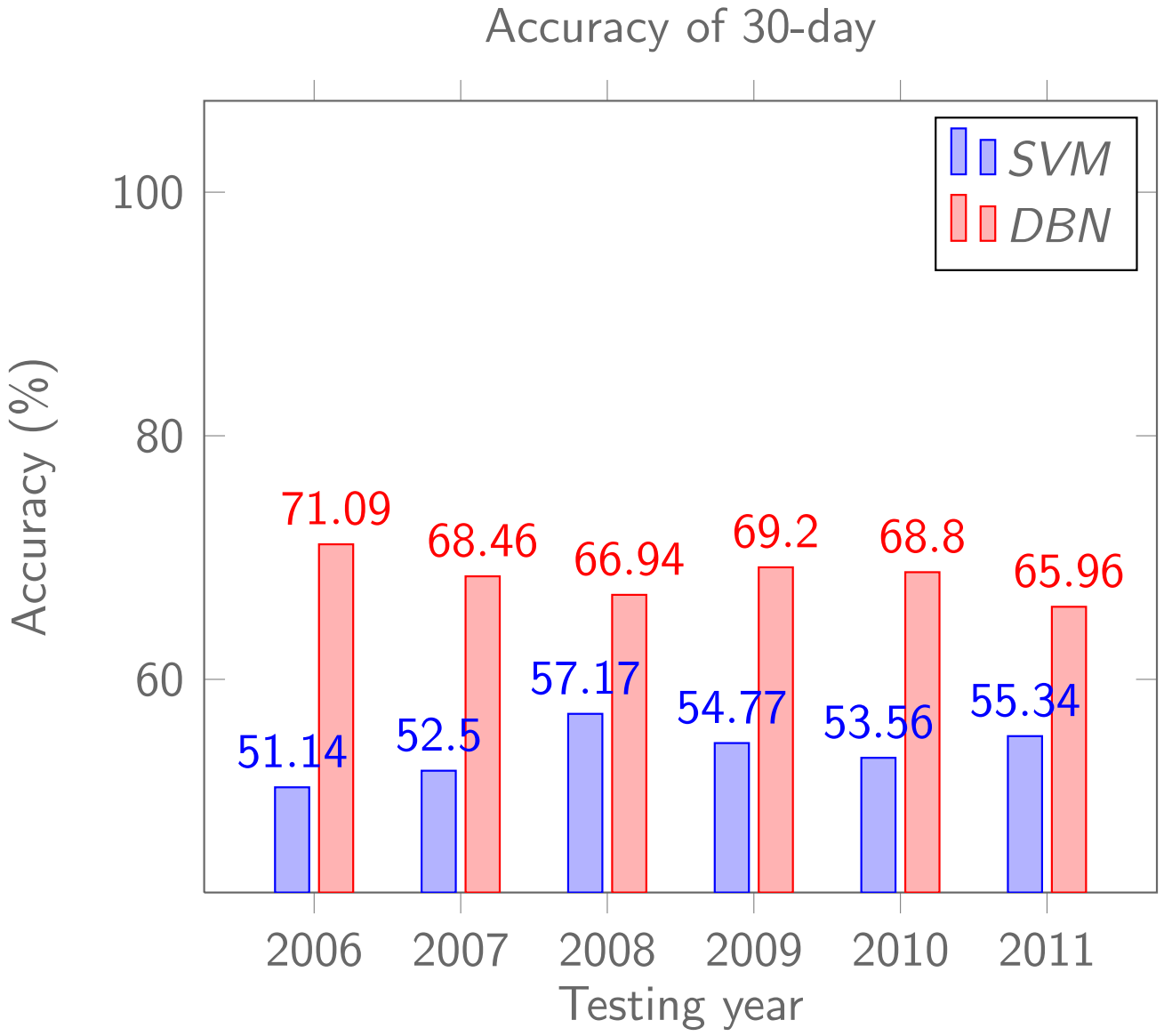
- Baseline (Classifier: Support Vector Classification via LIBSVM[3])
  - Features:
    1. 30-day: prior to default 5, 10, 15, 30-days average returns
    2. 180-day: prior to default 5, 10, 15, 30, 90, 180-days average returns
    3. 360-day: prior to default 5, 10, 15, 30, 90, 180, 360-days average returns
  - The training data is composed of the record in a five-year period, the following year of which is the testing data.
    - e.g., 2001-2005 for training and 2006 for testing.

[3] http://www.csie.ntu.edu.tw/~cjlin/libsvm/

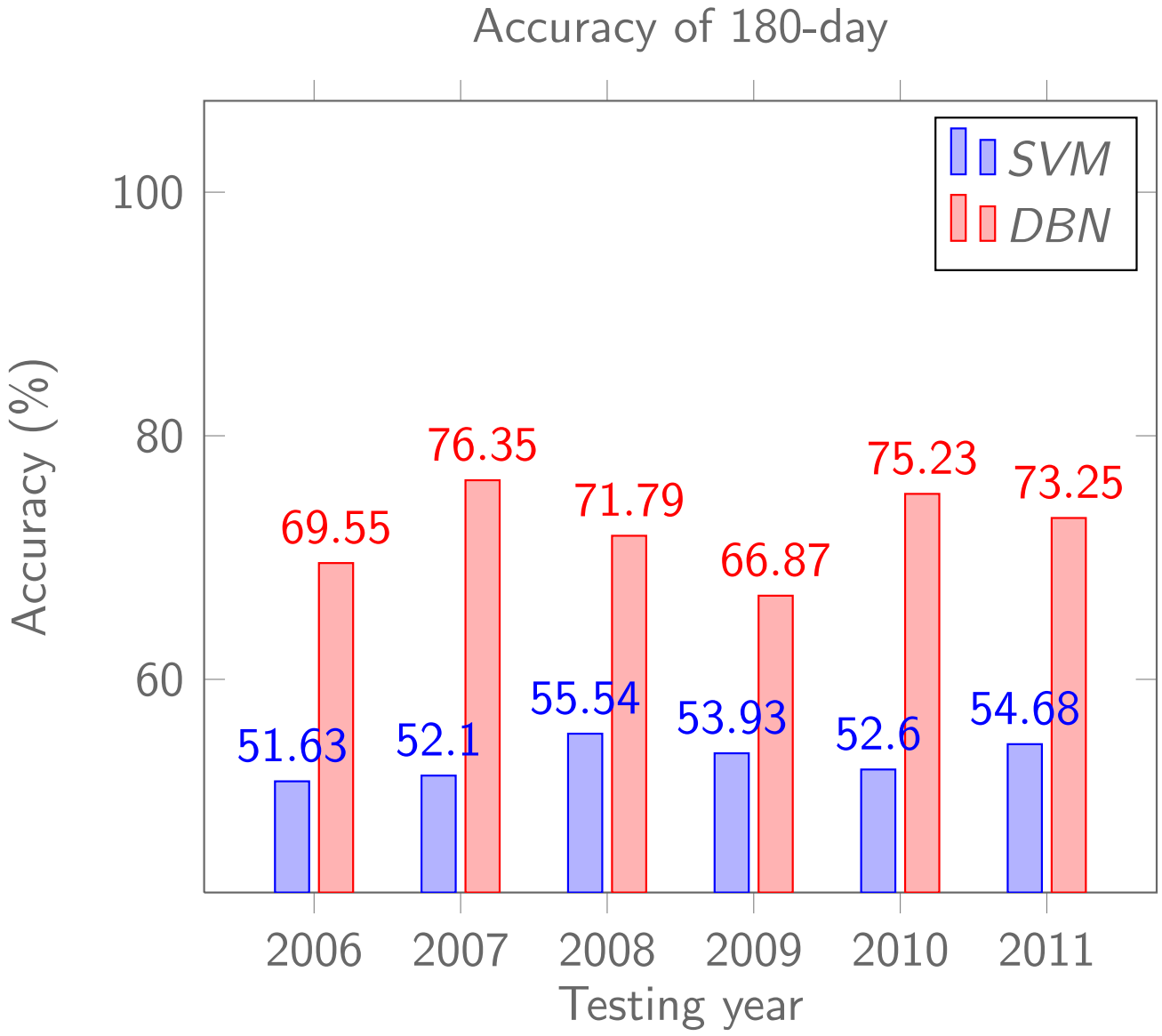
# Experimental Settings

- Our experiments (Classifier: DBN via theano)
  - The training data is composed of the record in a four-year period, the following year of which is the validation data, the next year is the testing data.
    - e.g., 2001-2004 for training, 2005 for validation, and 2006 for testing.

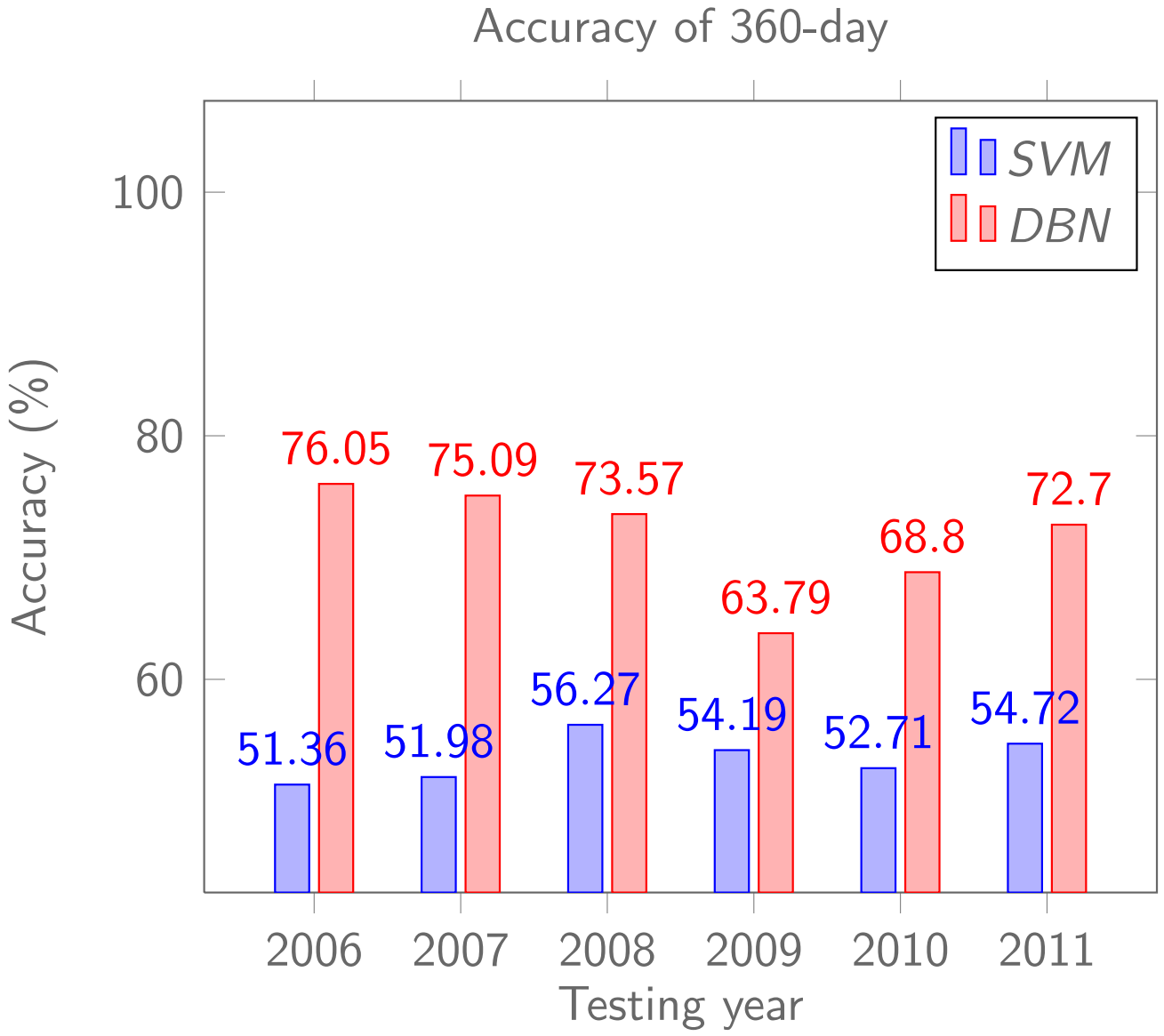
# Experiment Results



# Experiment Results



# Experiment Results



# Outline

# Conclusion

- This paper provides a new perspective on the default prediction problem using deep learning algorithms.
  - The representable factors of input data will no longer need to be explicitly extracted but can be implicitly learned by the learning algorithms.
  - We consider the stock returns of both default and solvent companies as input signals with graph representations, and use Deep Belief Networks to train the prediction models.
- In our experiments, the deep learning algorithm outperform better than traditionally machine learning algorithms.
- Future work: Identify and analyze the representation of the input signals.