Corporate Default Prediction via Deep Learning

Shu-Hao Yeh,\textsuperscript{1} Chuan-Ju Wang,\textsuperscript{1} Ming-Feng Tsai\textsuperscript{2}

\textsuperscript{1}University of Taipei, Taipei, Taiwan
\textsuperscript{2}National Chengchi University, Taipei, Taiwan

July 1, 2014
1 Introduction

2 Methodology

3 Experiments

4 Conclusion
Lehman Brothers

<table>
<thead>
<tr>
<th>Company</th>
<th>Filing date</th>
<th>Total Assets pre-filing</th>
<th>Assets adjusted to the year 2012</th>
<th>Filing court district</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lehman Brothers Holdings Inc.</td>
<td>2008-09-15</td>
<td>$639,063,000,800</td>
<td>$700 billion</td>
<td>NY-S</td>
</tr>
</tbody>
</table>

**Lehman Brothers’ share price**

Tuesday’s close: $7.79
Down 45%

Source: WSJ Market Data Group
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
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?
Deep learning

1. A machine learning method based on learning representations
   → Different concepts are learned from other concepts, with the more abstract, higher level concepts being learned from the lower level ones.
   → Deep learning helps to disentangle these abstractions and pick out which features are useful for learning.

   → Applied on computer vision, automatic speech recognition and natural language processing
Main contributions

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

\[-0.098684, -0.138686, 0.016949, \ldots, -0.365854, 0.076923\]
Main contributions

1. 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
Daily stock return

\[ r_t = \frac{S_t - S_{t-1}}{S_{t-1}} \]

\( S_{t-1} = \) stock price at day \( t - 1 \)

\( S_t = \) stock price at day \( t \)
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/
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

<table>
<thead>
<tr>
<th>PERMNO</th>
<th>Date</th>
<th>Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>10001</td>
<td>20030102</td>
<td>0.088287</td>
</tr>
<tr>
<td>10001</td>
<td>20030103</td>
<td>0.052500</td>
</tr>
<tr>
<td>10001</td>
<td>20030106</td>
<td>-0.019121</td>
</tr>
<tr>
<td>10001</td>
<td>20030107</td>
<td>0.004964</td>
</tr>
<tr>
<td>10001</td>
<td>20030108</td>
<td>0.006024</td>
</tr>
</tbody>
</table>

²https://wrds-web.wharton.upenn.edu/wrds/
## Dataset

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

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/
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

Accuracy of 30-day

Accuracy (%)

Testing year

- SVM
- DBN

2006: 51.14, 71.09
2007: 52.5, 68.46
2008: 57.17, 66.94
2009: 54.77, 69.2
2010: 53.56, 68.8
2011: 55.34, 65.96
Experiment Results

Accuracy of 180-day

<table>
<thead>
<tr>
<th>Testing year</th>
<th>SVM</th>
<th>DBN</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>51.63</td>
<td>69.55</td>
</tr>
<tr>
<td>2007</td>
<td>52.10</td>
<td>76.35</td>
</tr>
<tr>
<td>2008</td>
<td>55.54</td>
<td>71.79</td>
</tr>
<tr>
<td>2009</td>
<td>53.93</td>
<td>66.87</td>
</tr>
<tr>
<td>2010</td>
<td>52.60</td>
<td>75.23</td>
</tr>
<tr>
<td>2011</td>
<td>54.68</td>
<td>73.25</td>
</tr>
</tbody>
</table>
Experiment Results

Accuracy of 360-day

<table>
<thead>
<tr>
<th>Testing year</th>
<th>SVM</th>
<th>DBN</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>51.36</td>
<td>76.05</td>
</tr>
<tr>
<td>2007</td>
<td>51.98</td>
<td>75.09</td>
</tr>
<tr>
<td>2008</td>
<td>56.27</td>
<td>73.57</td>
</tr>
<tr>
<td>2009</td>
<td>54.19</td>
<td>63.79</td>
</tr>
<tr>
<td>2010</td>
<td>52.71</td>
<td>68.8</td>
</tr>
<tr>
<td>2011</td>
<td>54.72</td>
<td>72.7</td>
</tr>
</tbody>
</table>
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.