

FISH: A Financial Interactive System for Signal Highlighting

Ta-Wei Huang^{1*}, Jia-Huei Ju^{1*}, Yu-Shiang Huang¹, Cheng-Wei Lin¹, Yi-Shyuan Chiang², and Chuan-Ju Wang¹

¹Research Center for Information Technology Innovation, Academia Sinica

²Department of Computer Science, University of Illinois, Urbana-Champaign

Introduction

Overview and demonstration

Real-world scenarios

- Reviewing financial document requires lots of efforts.
- Human can hardly find out important information in such a short period of time.

Our goals

- Facilitate the reviewing process of financial documents analysis.
- Develop a BERT-based system to help users pinpoint signals in the tedious document.

- 2-min video: <https://www.youtube.com/watch?v=ZbvZQ09i6aw>



- Web-based online demonstration: <https://cfda.csie.org/FISH/>

Interface

Company metadata menu

The screenshot shows a web browser window with the URL `fish4heroku.herokuapp.com/report1`. The application header includes the logo 'FISH' and a navigation menu with three dropdown menus: 'Company List', 'Report Year', and 'Item', followed by an 'Anlyse' button. Below the header, the company name 'COHERENT INC' is displayed. A legend indicates segment types: Revised segments (orange dot), Highly similar segments (purple dot), and New segments (green dot).

The main content area is split into two columns. The left column is titled 'from 2012' and the right column is titled 'from 2011'. Both columns display a 'Business Overview' section with highlighted text segments. The 2012 version includes a green highlight for 'Fiscal years 2012, 2011 and 2010 included 52 weeks.' The 2011 version includes a purple highlight for 'Fiscal years 2011 and 2010 included 52 weeks; fiscal year 2009 included 53 weeks.'

Vertical scroll indicators are present on both sides of the text blocks, and a legend at the bottom right identifies the segment colors: Revised segments (orange), Highly similar segments (purple), and New segments (green).

Interface

Report section bookmarks

The screenshot shows a web browser window with the URL `fish4heroku.herokuapp.com/report1`. The application header includes the logo 'FISH' and navigation controls: 'Company List', 'Report Year', 'Item', and an 'Analyse' button.

ESTEE LAUDER COMPANIES INC

Legend: ● Revised segments ● Highly similar segments ● New segments

from 2013

1 Since certain promotional activities are a component of sales or cost of sales and the timing and level of promotions vary with our promotional calendar, we have experienced, and expect to continue to experience, fluctuations in the cost of sales percentage. In addition, future cost of sales mix may be impacted by the inclusion of potential new brands or channels of distribution that have margin and product cost structures different from those of our current mix of business.

1A

1B
2
3
4
5
6
7

Operating expenses as a percentage of net sales **decreased** to 65.1% as compared with 66.0% in the prior year. This improvement reflected a decrease in general and administrative costs as a percentage of net sales of approximately 50 basis points, a decrease in charges associated with restructuring activities of approximately 40 basis points and lower selling and shipping costs as a percentage of net sales of 10 basis points. Also included in this improvement was a

from 2012

1 Since certain promotional activities are a component of sales or cost of sales and the timing and level of promotions vary with our promotional calendar, we have experienced, and expect to continue to experience, fluctuations in the cost of sales percentage. In addition, future cost of sales mix may be impacted by the inclusion of potential new brands or channels of distribution that have margin and product cost structures different from those of our current mix of business.

1A

1B
2
3
4
5
6
7

Operating expenses as a percentage of net sales **increased** to 66.0% as compared with 65.6% in the prior year. This change reflected increased spending in advertising, merchandising and sampling costs in line with our strategy of approximately 80 basis points, higher costs related to stock-based compensation of approximately 20 basis points, an increase in general and administrative costs of approximately 10 basis points and higher charges associated with

Interface

Two-panel design for better comparisons

FISH

Company List Report Year Item Analyze

ESTEE LAUDER COMPANIES INC

from 2013

1 The lower net sales in Russia primarily reflected destocking associated with ongoing challenges with a certain customer. The overall change in Europe, the Middle East Africa net sales was inclusive of unfavorable exchange rates due to the strengthening of the U.S. dollar against most currencies in this region of approximately \$75 million. Excluding the impact of foreign currency translation, net sales in Europe, the Middle East Africa increased 6%.

1A

1B

2

3

4

5

6

7

Net sales in Asia/Pacific increased 5%, or \$110.2 million, to \$2,121.6 million, primarily reflecting growth in our sales of skin care products, in line with our strategy. We increased sales by approximately \$160 million in China and Hong Kong. Net sales in China benefited from expanded distribution. While we gained share in the prestige business in China, we are cautious that a slowing of the future growth trend of the Chinese economy may temper our retail sales growth, including that of our travel retail business. Higher sales in Hong Kong were

from 2012

1 favorable effect of exchange rates of approximately 10 basis points. These improvements were partially offset by an increase in obsolescence charges of approximately 10 basis points.

1A Since certain promotional activities are a component of sales or cost of sales and the timing and level of promotions vary with our promotional calendar, we have experienced, and expect to continue to experience, fluctuations in the cost of sales percentage. In addition, future cost of sales mix may be impacted by the inclusion of potential new brands or channels of distribution that have margin and product cost structures different from those of our current mix of business.

1B

2

3

4

5

6

7

Operating expenses as a percentage of net sales increased to 66.0% as compared with 65.6% in the prior year. This change reflected increased spending in advertising, merchandising and

● Revised segments ● Highly similar segments ● New segments

Main Features

Sentence-level highlighting

from 2012

1 amounts do not necessarily represent future cash requirements.

1A The amount of collateral obtained, if deemed necessary by the Heartland banks upon extension of credit, is based upon managements credit evaluation of the counterparty.

1B Collateral held varies but may include accounts receivable, inventory, property, plant and equipment and income-producing commercial properties.

Standby letters of credit and financial guarantees written are conditional commitments issued by the Heartland banks to guarantee the performance of a customer to a third party. Those guarantees are primarily issued to support public and private borrowing arrangements. The credit risk involved in issuing letters of credit is essentially the same as that involved in extending loan facilities to customers. At December 31, 2011, and December 31, 2010, commitments to extend credit aggregated \$765.8 million and \$623.2 million, and standby letters of credit aggregated \$49.1 million and \$48.7 million, respectively.

9A The following table summarizes our significant contractual obligations and other commitments as of December 31, 2011:

from 2011

1 The Bank Subsidiaries evaluate each customer s creditworthiness on a case-by-case basis. The amount of collateral obtained, if deemed necessary by the Bank Subsidiaries upon extension of credit, is based upon management s credit evaluation of the counterparty. Collateral held varies but may include accounts receivable, inventory, property, plant and equipment and in other public income producing properties.

Standby letters of credit and financial guarantees written are conditional commitments issued by the Bank Subsidiaries to guarantee the performance of a customer to a third party. Those guarantees are primarily issued to support public and private borrowing arrangements. The credit risk involved in issuing letters of credit is essentially the same as that involved in extending loan facilities to customers.

The following table summarizes our significant contractual obligations and other commitments as of December 31, 2010:

On a consolidated basis, we maintain a large balance of short-term securities that, when

Highly similar

Revised

New

Main Features

Token-level highlighting

from 2012

systems and mission support services.

Our transportation systems business is the leading provider of automated revenue collection systems and services worldwide. We provide complete turnkey solutions. Our equipment includes contactless smart card readers, passenger gates, central computer systems, and ticket vending machines for mass transit networks, including rail systems, buses, and parking applications. Our services include customer support, network and web operations, payment media

management, distribution channel management, business and marketing support, financial clearing and settlement, and outsourced asset operations and maintenance.

Our defense systems business includes training systems, communications, cyber security and asset tracking. We are a leading provider of customized military

from 2011

and mission support services.

Our transportation systems business is the leading provider of automated revenue collection systems and services worldwide. We provide complete solutions. Our equipment includes contactless smart card readers, passenger gates, central computer systems, and ticket vending machines for mass transit networks, including rail systems, buses, and parking applications. Our services include customer support, network and web operations, payment media

management, distribution channel management, business and marketing support, financial clearing and settlement, and outsourced asset operations and maintenance.

Our defense systems business includes training systems, communications, cyber security and asset tracking. We are a leading provider of customized military

Predicted
word importance

The Pipeline Approach

Segment classifier

To narrow considerations of analyzing all the segments in a financial report, we classify each of them into three types

- **New / Highly-similar / Revised segments**

Based on two **similarity scores** to discriminate the types

- ROUGE-2 (syntactic)
- SBERT cosine (semantic)

$$s_t^i \text{ type} = \begin{cases} 1 & \text{if } \max(\{\phi_{\text{Rouge}}(s_t^i, s_{t-1}^j) | s_{t-1}^j \in \mathcal{S}_{t-1}\}) < \tau \\ 2 & \text{if } \max(\{\phi_{\text{Rouge}}(s_t^i, s_{t-1}^j) | s_{t-1}^j \in \mathcal{S}_{t-1}\}) \geq \tau \\ & \text{AND } \phi_{\text{BERT}}(s_t^i, s_{t-1}^{j*}) \geq \epsilon, \\ 3 & \text{if } \max(\{\phi_{\text{Rouge}}(s_t^i, s_{t-1}^j) | s_{t-1}^j \in \mathcal{S}_{t-1}\}) \geq \tau \\ & \text{AND } \phi_{\text{BERT}}(s_t^i, s_{t-1}^{j*}) < \epsilon, \end{cases} \quad (1)$$

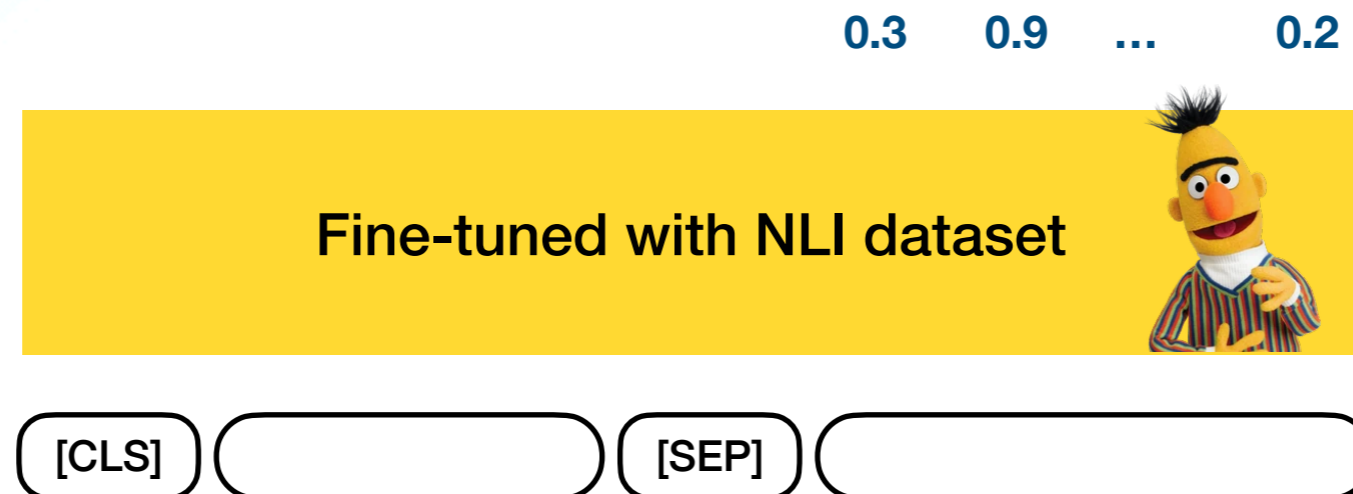
The Pipeline Approach

Segment highlighter

To further pinpoint particular signals in a fine-grained manner, we leverage the contextualized token representation of BERT:

$$h_{rj^*} = \text{BERT}([\text{CLS}]s_{t-1}^{j^*}[\text{SEP}]s_t^r).$$

And perform **word importance prediction** by regarding it as a sequence labeling task:



Thank you!

Are there any questions you'd like to ask?

Ta-Wei Huang, a0917589225@gmail.com

Jia-Huei Ju, jhjoo@citi.sinica.edu.tw

Chuan-Ju Wang, cjwang@citi.sinica.edu.tw