Sentiment Analysis for Credit Risk and Portfolio Construction

Christina Erlwein-Sayer
Xiang Yu
Overview

- Eurostars project SENRISK
  - Aim and Purpose
  - Partners and Product
    - News-enhanced bond modelling and forecasting
      - Use and system requirements, data management and processing
      - News and sentiment evaluation
      - Development of sentiment-enhanced key figures
  - Sentiment Analysis for Portfolio Construction

OptiRisk Systems

Eurostars
SENISK - Aim and purpose

Project aim:
- Development of an automated credit risk assessment tool
- Innovative Decision Support System for risk assessment of Fixed Income products incorporating news sentiments

Purpose:
- Valuation of sovereign and corporate bonds incorporating news-based information from the market.
- Predictive risk models shall be enhanced through information such as interest rates and macroeconomic factors as well as sentiments from firm-specific or macroeconomic news and social media sentiment.
SENISK - Consortium

OptiRisk Systems

Fraunhofer ITWM

Partners

PS Quant

ACATIS

OptiRisk Systems
eurostars™
The sentiment of macroeconomic news and social media items are combined with historical time series to construct a generic credit risk assessment tool.

Monitoring sovereign and corporate bonds is made more informative and efficient, the creditworthiness of sovereigns is assessed.

With our tool, all investors shall be able to handle the vast amount of information on European sovereigns.

Combined with the analysis of sentiment, accurate and relevant information is included in investment decisions.
DSS Platform

Inputs
- Sovereign & corporate bond market data
- Social media scores
- Market & macroeconomic news scores
- Fundamental reports

Risk assessment tool
- Filter & process data and scores
- Create scores
- Database (market data & scores (social media, market and macro news, fundamental))
- Enhanced risk assessment model

Data management

Business logic

Valuation

Front end
- Graphical user interface

OptiRisk Systems

Eurostars™
User and System requirements

- User requirements have been established through interviews with expert users and feedback on our proposed credit risk assessment tool.

- The data requirements have been specified in terms of markets and instruments needed as well as in terms of solutions for news and sentiment sources.

  - Markets: European mainland countries, UK and US
  - Short and long term bonds
  - Sovereign and corporate bonds
  - Source: Datascope (Thomson Reuters)
User and System requirements

- News and sentiment sources:
  - Sentiment data for relevant instruments
  - Macroeconomic news sentiment
    - Source: Ravenpack
    - Relevance and Sentiment considered, economic news item concerning the issuing country
  - Social media sentiment
    - Source: Stockpulse
Bond data: term structure specification

- Calculate term structure from bonds of specific country
  - Through the Svensson model calculate spread with (AAA) Eurobond from ECB
- Example: Greece
  - April 2016-May 2017
  - Short and long term maturity
Data management and (pre-) processing

- Input: News and social media sentiment score data, macroeconomic and market data
- Information is retrieved via APIs. After aggregating, filtering and processing, data is stored in the database.
- MySQL Server is the Database Management System, multi-user access to databases.
News and sentiment evaluation: news and social media contents

Ontology definition and Event research: macroeconomic news

Example: Macronews Germany (Ravenpack)

<table>
<thead>
<tr>
<th>year</th>
<th>date</th>
<th>hour</th>
<th>relevance</th>
<th>topic</th>
<th>groupCategory</th>
<th>category</th>
<th>typeEvent</th>
<th>factLevel</th>
<th>newsType</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>20070102</td>
<td>2600</td>
<td>7</td>
<td>economy</td>
<td>balance-of-payments</td>
<td>exports-up-exporter</td>
<td>exports</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>2800</td>
<td>30</td>
<td>economy</td>
<td>consumption</td>
<td>consumer-confidence-down</td>
<td>consumer-confidence</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>60213</td>
<td>1</td>
<td>economy</td>
<td>employment</td>
<td>unemployment-guidance</td>
<td>unemployment-guidance</td>
<td>forecast</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>60309</td>
<td>1</td>
<td>economy</td>
<td>employment</td>
<td>unemployment-guidance</td>
<td>unemployment-guidance</td>
<td>forecast</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>64556</td>
<td>1</td>
<td>economy</td>
<td>employment</td>
<td>unemployment-guidance</td>
<td>unemployment-guidance</td>
<td>forecast</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>73700</td>
<td>100</td>
<td>economy</td>
<td>employment</td>
<td>employment-up</td>
<td>employment</td>
<td>fact</td>
<td>NEWS-FLASH</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>92000</td>
<td>13</td>
<td>economy</td>
<td>consumption</td>
<td>consumer-confidence-down</td>
<td>consumer-confidence</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>92000</td>
<td>29</td>
<td>economy</td>
<td>consumption</td>
<td>consumer-confidence-down</td>
<td>consumer-confidence</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>92500</td>
<td>2</td>
<td>economy</td>
<td>employment</td>
<td>unemployment-down</td>
<td>unemployment</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>123800</td>
<td>1</td>
<td>economy</td>
<td>products-services</td>
<td>market-exit-location</td>
<td>market-exit</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>125800</td>
<td>3</td>
<td>economy</td>
<td>consumption</td>
<td>retail-sales</td>
<td>retail-sales</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>141900</td>
<td>38</td>
<td>economy</td>
<td>consumption</td>
<td>consumer-confidence-down</td>
<td>consumer-confidence</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>141900</td>
<td>38</td>
<td>economy</td>
<td>consumption</td>
<td>retail-sales-up</td>
<td>retail-sales-up</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
<tr>
<td>2007</td>
<td>20070102</td>
<td>142200</td>
<td>100</td>
<td>business</td>
<td>labor-issues</td>
<td>layoffs-location</td>
<td>layoffs</td>
<td>fact</td>
<td>FULL-ARTICLE</td>
</tr>
</tbody>
</table>

OptiRisk Systems

Eurostars™

Sentiment Analysis in Finance
News and sentiment evaluation: news and social media contents

- Sentiment classification, trustworthiness and relevance analysis for social media sentiment
- StockPulse sentiment for EuroStoxx 50 companies
- Daily sentiment is analysed and a relationship to bonds investigated
Development of sentiment-enhanced key figures

- Analysis of country-specific macroeconomic news sentiment as well as firm-specific sentiment
- Aggregation of intra-day sentiment to daily sentiment and impact values
- Distinguish between positive and negative news, creation of daily impact scores with decay
- Analysis of volume of news
Sentiment-enhanced key figures

Number of positive and negative news (2014 - 2017)

- Train collision
- Decline in GDP
- Unemployment on record low
- Slowdown in German economy
Sentiment-enhanced key figures

Spread and Sentiment

- spread
- positive news
- negative news
Sentiment-enhanced key figures

Spread and Impact

- spread
- positive news
- negative news
- all news

OptiRisk Systems

eurostars™
Sentiment-enhanced key figures

Spread and Sentiment

Date

spread — positive news — negative news — all news
Development of sentiment-enhanced key figures

Effect research: inference of news events and bond prices

- Sentiment impact evaluation
- Integration of sentiments to spread and volatility prediction of Fixed Income products
- Integration of sentiments to default probabilities and rating migration
Development of sentiment-enhanced key figures

Modelling bond closing yield series:

- Create spread with benchmark (AAA ECB rate)
- Correlation between sentiment series and spread series low but significant
- Granger causality between spread and sentiment series
- Linear regression analysis points to most relevant regressors being “number of all news”, “positive sentiment”, negative sentiment
- Modelling closing yields through ARIMA model with external regressor
Development of sentiment-enhanced key figures

- Model order: ARIMA(2,1,2)
- Fit and Forecast without external regressor and with positive/negative sentiment scores, impact values and news counts
- Fit and Forecast are improved when including external regressors (w.r.t. various error measures)
- Choice of best regressor seems to be dependent on yield series, however, negative series have a higher impact than positive ones
Development of sentiment-enhanced key figures

1st example: 30-year, issue date: 5. January 2007
## Development of sentiment-enhanced key figures

### Error measures: Fit

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAE</th>
<th>MPE</th>
<th>MAPE</th>
<th>MASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without regressor</td>
<td>0.03740</td>
<td>0.02443</td>
<td>-1.57051</td>
<td>22.1638</td>
<td>0.92294</td>
</tr>
<tr>
<td>Nr of all and pos news, all and pos impact</td>
<td>0.03734</td>
<td>0.02436</td>
<td>-1.82358</td>
<td>22.0655</td>
<td>0.92043</td>
</tr>
<tr>
<td>Nr of all and negative news, all and neg sentiment</td>
<td>0.03732</td>
<td>0.02440</td>
<td>-2.6035</td>
<td>22.9045</td>
<td>0.92179</td>
</tr>
<tr>
<td>All, pos and neg sentiment</td>
<td>0.03733</td>
<td>0.02434</td>
<td>-2.2642</td>
<td>22.5730</td>
<td><strong>0.91962</strong></td>
</tr>
<tr>
<td>Pos and Neg Impact</td>
<td>0.03735</td>
<td>0.02434</td>
<td>-2.4722</td>
<td>22.9108</td>
<td>0.91962</td>
</tr>
<tr>
<td>Pos Sentiment</td>
<td>0.03737</td>
<td>0.02438</td>
<td>-1.2208</td>
<td><strong>21.6723</strong></td>
<td>0.92125</td>
</tr>
<tr>
<td>All Impact</td>
<td>0.03738</td>
<td>0.024397</td>
<td>-1.7513</td>
<td>22.1099</td>
<td>0.92174</td>
</tr>
<tr>
<td>Nr of all news</td>
<td>0.03741</td>
<td>0.024434</td>
<td>-1.6730</td>
<td>22.2419</td>
<td>0.92313</td>
</tr>
</tbody>
</table>
## Development of sentiment-enhanced key figures

### Error measures: Forecast

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAE</th>
<th>MPE</th>
<th>MAPE</th>
<th>MAESE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without regressor</td>
<td>0.02720</td>
<td>0.02051</td>
<td>10.0162</td>
<td>25.7071</td>
<td>0.8407</td>
</tr>
<tr>
<td>Nr of all and pos news, all and pos impact</td>
<td>0.02724</td>
<td>0.02070</td>
<td>9.9764</td>
<td>25.7434</td>
<td>0.84857</td>
</tr>
<tr>
<td>Nr of all and negative news, all and neg sentiment</td>
<td>0.02758</td>
<td>0.02105</td>
<td>10.1721</td>
<td>26.2114</td>
<td>0.86287</td>
</tr>
<tr>
<td>All, pos and neg sentiment</td>
<td>0.02729</td>
<td>0.02062</td>
<td>10.05487</td>
<td>25.7888</td>
<td>0.84543</td>
</tr>
<tr>
<td>Pos and Neg Impact</td>
<td>0.02728</td>
<td>0.02063</td>
<td>10.1321</td>
<td>25.8617</td>
<td>0.84551</td>
</tr>
<tr>
<td>Pos Sentiment</td>
<td>0.02714</td>
<td>0.02044</td>
<td>9.9308</td>
<td><strong>25.5419</strong></td>
<td><strong>0.8380</strong></td>
</tr>
<tr>
<td>All Impact</td>
<td>0.02722</td>
<td>0.02063</td>
<td><strong>9.8114</strong></td>
<td>25.5686</td>
<td>0.8457</td>
</tr>
<tr>
<td>Nr of all news</td>
<td>0.02716</td>
<td>0.02051</td>
<td>9.9525</td>
<td>25.6486</td>
<td>0.84077</td>
</tr>
</tbody>
</table>
Development of sentiment-enhanced key figures

2nd example: 10-year bond, issued on 12th April 2012
# Development of sentiment-enhanced key figures

## Error measures: Fit

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAE</th>
<th>MPE</th>
<th>MAPE</th>
<th>MASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without regressor</td>
<td>0.023123</td>
<td>0.015898</td>
<td>-2.7025</td>
<td>9.9953</td>
<td>0.8868</td>
</tr>
<tr>
<td>Nr of all and pos news, all and pos impact</td>
<td>0.023110</td>
<td>0.015885</td>
<td>-2.6970</td>
<td>9.9788</td>
<td>0.8861</td>
</tr>
<tr>
<td>Nr of all and negative news, all and neg sentiment</td>
<td><strong>0.023107</strong></td>
<td>0.015876</td>
<td>-2.7075</td>
<td>9.9844</td>
<td>0.8856</td>
</tr>
<tr>
<td>All, pos and neg sentiment</td>
<td>0.023110</td>
<td><strong>0.015872</strong></td>
<td>-2.7139</td>
<td>9.9933</td>
<td><strong>0.8854</strong></td>
</tr>
<tr>
<td>Pos and Neg Impact</td>
<td>0.023117</td>
<td>0.015889</td>
<td>-2.6991</td>
<td>9.9905</td>
<td>0.8863</td>
</tr>
<tr>
<td>Pos Sentiment</td>
<td>0.023121</td>
<td>0.015891</td>
<td>-2.7045</td>
<td>9.9894</td>
<td>0.8865</td>
</tr>
<tr>
<td>All Impact</td>
<td>0.023118</td>
<td>0.015887</td>
<td><strong>-2.6832</strong></td>
<td>9.9877</td>
<td>0.8862</td>
</tr>
<tr>
<td>Nr of all news</td>
<td>0.023121</td>
<td>0.015899</td>
<td>-2.6972</td>
<td>9.9854</td>
<td>0.8869</td>
</tr>
</tbody>
</table>
## Development of sentiment-enhanced key figures

### Error measures: Forecast

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
<th>MAE</th>
<th>MPE</th>
<th>MAPE</th>
<th>MASE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without regressor</td>
<td>0.018507</td>
<td>0.013773</td>
<td>-14.406</td>
<td>53.788</td>
<td>0.9315</td>
</tr>
<tr>
<td>Nr of all and pos news, all and pos impact</td>
<td>0.018404</td>
<td>0.013699</td>
<td>-14.108</td>
<td>53.577</td>
<td>0.82930</td>
</tr>
<tr>
<td>Nr of all and negative news, all and neg sentiment</td>
<td>0.018466</td>
<td>0.013737</td>
<td>-14.406</td>
<td>53.579</td>
<td>0.83160</td>
</tr>
<tr>
<td>All, pos and neg sentiment</td>
<td>0.018462</td>
<td>0.013721</td>
<td>-14.420</td>
<td>53.448</td>
<td>0.83075</td>
</tr>
<tr>
<td>Pos and Neg Impact</td>
<td>0.018447</td>
<td>0.013719</td>
<td>-14.278</td>
<td>53.616</td>
<td>0.83051</td>
</tr>
<tr>
<td>Pos Sentiment</td>
<td>0.018463</td>
<td>0.013728</td>
<td>-14.312</td>
<td>53.642</td>
<td>0.83101</td>
</tr>
<tr>
<td>All Impact</td>
<td>0.018470</td>
<td>0.013747</td>
<td>-14.309</td>
<td>53.636</td>
<td>0.83216</td>
</tr>
<tr>
<td>Nr of all news</td>
<td>0.018490</td>
<td>0.013758</td>
<td>-14.348</td>
<td>53.836</td>
<td>0.83285</td>
</tr>
</tbody>
</table>
Further steps

- Development of news-enhanced risk control decision model for Fixed Income Products
- Decision Support System: integration and system validation

Please visit our project page [www.senrisk.eu](http://www.senrisk.eu) for more information.
Sentiment Analysis for Portfolio Construction

- News Sentiment is utilised to
  - Improve scenario generation for portfolio construction
  - Choose suitable asset universe prior to run portfolio optimisation

- Sentiment input: firm-specific impact time series
  - Impact series is built w.r.t. time span and market place, decaying importance of news
Sentiment Analysis for Portfolio construction

Portfolio performance of Second-order Stochastic Dominance portfolio with news sentiment

Reference:
Thank you very much for your attention!

Any questions?