Motivation

Humor detection systems often use completely different texts as non-jokes, e.g. proverbs and news.

We use text imitation algorithms for generating non-jokes from jokes, increasing the challenge of humor detection.

Evaluation on different common models shows that only RobBERT is able to distinguish jokes from generated non-jokes.

Dynamic Templates algorithm

Automatically creates templates for content from base texts Used here to generate non-jokes from jokes.

Base text words:

Two fish are in a tank. Says one to the other: "Do you know how to drive this thing?"

Context text words:

A horse walks into a bar, says one of the men: why the long face?

Replace matching POS words, starting with lowest frequency b-words until minimum number of replacements is reached

Non-Joke Examples

RobBERT

Het is groen en het is een mummie? Kermit de Waterkant

"Ober, kunt u die schrik uit mijn politieman halen? Want ik eet liever alleen."

Hoe heet de vrouw van Sinterklaas? Keukentafel.

"Twee tanden zwemmen in de zee en ze zien een stamgast op een stamgast. De ene raad zegt tegen de andere raad: 'Hé kijk! Ons eten op een bord!'"

Data

3235 Dutch Jokes from KidsWeek.nl, DeBesteMoppen.nl and LachJeKrom.com

1887 Dutch Proverbs from Wikipedia

3235 News Headlines randomly selected from 100K Dutch news headlines dataset

3235 Generated Non-Jokes, generated using jokes dataset as input for the Dynamic Template algorithm to generate non-sensical joke-like texts

Evaluation

Binary classification of jokes versus texts from other domains

<table>
<thead>
<tr>
<th></th>
<th>Naive Bayes</th>
<th>LSTM</th>
<th>CNN</th>
<th>RobBERT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jokes vs News</td>
<td>51%</td>
<td>94%</td>
<td>94%</td>
<td>99%</td>
</tr>
<tr>
<td>Jokes vs Proverbs</td>
<td>60%</td>
<td>94%</td>
<td>94%</td>
<td>96%</td>
</tr>
<tr>
<td>Jokes vs Generated Non-Jokes</td>
<td>50%</td>
<td>47%</td>
<td>47%</td>
<td>89%</td>
</tr>
</tbody>
</table>

Findings: While models like LSTMs and CNNs are able to distinguish jokes from news and proverbs, only RobBERT is able to distinguish jokes from generated non-jokes, indicating that our new dataset is vastly more challenging.

Conclusion

Generating non-jokes from jokes creates a more challenging dataset, showing the power of transformer for humor detection.

Code & models: https://github.com/twinters/dutch-humor-detection