Fake or Real? A Study of Arabic Satirical Fake News

Hadeel Saadany, Emad Mohamed, Constantin Orasan

{h.a.saadany, e.mohamed2} @wlv.ac.uk, c.Orasan@surrey.ac.uk
Outline:

Part One:
- Is satirical news harmful?
- Satirical fake news vs. propaganda and hoax.
- Hypothesis:
  *Linguistic features per se suffice to detect satirical fake news with high accuracy*

Part Two:
- ML models used for detecting Arabic Satirical Fake news.
- ML satirical fake news detection with accuracy = 98.56%
Part One
Is Satirical Fake News humorous or harmful?

Mexican tourists killed by Egyptian security forces

#thankyou_Henrico_Iglassios
Taking Satire Seriously

• The Onion
• The Beaverton
• Etc.

"Congress Passes Americans With No Abilities Act"

“FIFA Frantically Announces 2015 Summer World Cup In United States”
Automatic fake news detection: Satire vs Hoax and Propaganda

Features:
- Propaganda and hoax: meta-data
  - The source
  - Reputation of the publisher
  - History of truthfulness
  - Linguistic

- Satirical: Linguistic
- Hoax and propaganda: low accuracy (20% - 70%)
- Satirical: high accuracy (85%-99%)
Part Two
Arabic Satirical Fake News Detection:

Data compiling:

- Fake dataset: (3185 articles)
- Arabic satirical online news website (Al Hudood, Mexican Ahram)
- Fact dataset: (3710 articles)
- BBC-Arabic, CNN-Arabic, Al-Jazeera news articles

Linguistic-Statistical Analysis

- Journalistic Register Measure
- Stated the spokesman \( \text{قال الناطق باسم} \), Reported by.. \( \text{تقرير أعده} \)

\[
J = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} I_C(\omega)
\]

\[
I_C(\omega) = \begin{cases} 
1 & \omega \in C \\
0 & \omega \notin C 
\end{cases}
\]

Figure 1: Probability Density Function of Journalistic Register Measure
Sentiment Intensity Measure

• Satire is created by emotional imbalance
• Statistically significant (statistic 3.27, p-value 0.001)

Subjectivity Measure

• The ratio of first-person plural verbs in a satirical fake article to a real news article was 2.2 to 1.
Classification Models and Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Test Set</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acc</td>
<td>P</td>
<td>R</td>
<td>F</td>
</tr>
<tr>
<td>Multinomial NB (Count Vectors)</td>
<td>96.23</td>
<td>96.47</td>
<td>96.23</td>
<td>96.24</td>
</tr>
<tr>
<td>Multinomial NB (Word-Level TF-IDF)</td>
<td>94.63</td>
<td>94.71</td>
<td>94.64</td>
<td>94.64</td>
</tr>
<tr>
<td>Multinomial NB (N-Grams)</td>
<td>86.08</td>
<td>88.03</td>
<td>86.09</td>
<td>86.22</td>
</tr>
<tr>
<td>Multinomial NB (CharLevel)</td>
<td>91.73</td>
<td>92.00</td>
<td>91.74</td>
<td>91.75</td>
</tr>
<tr>
<td>Multinomial NB (Count Vectors segmented)</td>
<td>95.79</td>
<td>95.97</td>
<td>95.80</td>
<td>95.81</td>
</tr>
<tr>
<td>Multinomial NB (Word-Level TF-IDF segmented)</td>
<td>94.49</td>
<td>94.65</td>
<td>94.49</td>
<td>94.51</td>
</tr>
<tr>
<td>XGBoost with Count Vectors</td>
<td>96.81</td>
<td>96.81</td>
<td>96.81</td>
<td>96.81</td>
</tr>
<tr>
<td>CNN with pre-trained word embeddings</td>
<td>98.59</td>
<td>98.49</td>
<td>98.61</td>
<td>98.49</td>
</tr>
</tbody>
</table>

Table 1: Summary of Accuracy (Acc), Precision (P), Recall (R) and F1 score (F) of classification models.
Error Analysis:

Confusion matrix for baseline model
• Why a fake article is classified as real?

1. A perfect parody of the real news with a proliferation of journalistic clichés (15 phrases).
2. Written in MSA (Modern Standard Arabic), no slang.
3. Strong sentiment polarity but were not captured by the classifier (mostly coined hyperbole’s specific of the author’s style)
Conclusion:

• A novel problem for Arabic Computational Linguistics
• Fake News detection accuracy is not the same for all types.
• Satirical fake news detection can be achieved by AI solutions.
Thank you