Detecting Protagonists in German Plays around 1800 as a Classification Task

Nils Reiter, Benjamin Krautter, Janis Pagel, Marcus Willand
Disclaimer

- We have progressed in our work
- We will present the current state of our research
- **Not** only what is in the submitted paper
- Updated results
- Talk includes analysis on dramas from 1700 to 1900
Quantitative Drama Analytics (QuaDramA)

- Cooperation between German literary studies and computational linguistics
Quantitative Drama Analytics (QuaDramA)

- Cooperation between German literary studies and computational linguistics
- Analyse (German) dramatic texts computationally
Quantitative Drama Analytics (QuaDramA)

- Cooperation between German literary studies and computational linguistics
- Analyse (German) dramatic texts computationally
- Historical perspective
Quantitative Drama Analytics (QuaDramA)

• Cooperation between German literary studies and computational linguistics
• Analyse (German) dramatic texts computationally
• Historical perspective
• Investigate character types
Quantitative Drama Analytics (QuaDramA)

- Cooperation between German literary studies and computational linguistics
- Analyse (German) dramatic texts computationally
- Historical perspective
- Investigate character types
- Coreference resolution for dramatic texts
Introduction
What is a Drama?

- Use of action and dialogue
What is a Drama?

- Use of action and dialogue
- Typically divided into acts and scenes
What is a Drama?

- Use of action and dialogue
- Typically divided into acts and scenes
- Typically designed to be performed (on stage)
What is a Drama?

- Use of action and dialogue
- Typically divided into acts and scenes
- Typically designed to be performed (on stage)
- Stage directions
What is a Drama?

- Use of action and dialogue
- Typically divided into acts and scenes
- Typically designed to be performed (on stage)
- Stage directions
- Cast
What is a Drama?

- Use of action and dialogue
- Typically divided into acts and scenes
- Typically designed to be performed (on stage)
- Stage directions
- Cast
- Dramatic conflict
Fünfter Auftritt

*Der Prinz. Emilia. Marinelli.*

**DER PRINZ.**
Wo ist sie? wo? -
Wir suchen Sie überall, schönstes Fräulein.
- Sie sind doch wohl?
- Nun so ist alles wohl!
Der Graf, Ihre Mutter, -

**EMILIA.**
Ah, gnädigster Herr!
wo sind sie?
Wo ist meine Mutter?
Fünfter Auftritt

Der Prinz. Emilia. Marinelli.

DER PRINZ.
Wo ist sie? wo? -
Wir suchen Sie überall, schönstes Fräulein.
- Sie sind doch wohl?
- Nun so ist alles wohl!
Der Graf, Ihre Mutter, -

EMILIA.
Ah, gnädigster Herr!
wo sind sie?
Wo ist meine Mutter?

<div type="scene">
  <div>
    <desc>
      <title>5. Auftritt</title>
    </desc>
  </div>
  <div type="text">
    <div type="h4">
      <head>Fünfter Auftritt</head>
      <stage>
        <hi>Der Prinz. Emilia. Marinelli.</hi>
      </stage>
      <sp who="#der_prinz">
        <speaker>DER PRINZ.</speaker>
        <p>Wo ist sie? wo? -
           Wir suchen Sie überall,
           schönstes Fräulein. -
           Sie sind doch wohl? -
           Nun so ist alles wohl!
           Der Graf, Ihre Mutter, -</p>
      </sp>
      <sp who="#emilia">
        <speaker>EMILIA.</speaker>
        <l>Ah, gnädigster Herr!
           wo sind sie?
           Wo ist meine Mutter?</l>
      </sp>
    </div>
  </div>
</div>
Goals

- Classify all figures in play regarding the classes: Protagonist - Not Protagonist
Goals

- Classify all figures in play regarding the classes: Protagonist - Not Protagonist
- Analyse results w.r.t. literary interpretation
Definition of Being a Protagonist

- Difficult from theoretical point of view
Definition of Being a Protagonist

- Difficult from theoretical point of view
- We settled on:

<table>
<thead>
<tr>
<th>Protagonist</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Causes or solves the central dramatic conflict</td>
</tr>
<tr>
<td>- This can happen actively or passively</td>
</tr>
</tbody>
</table>
Definition of Being a Protagonist

- Difficult from theoretical point of view
- We settled on:

<table>
<thead>
<tr>
<th>Protagonist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Causes or solves the central dramatic conflict</td>
</tr>
<tr>
<td>This can happen actively or passively</td>
</tr>
</tbody>
</table>

- From this follows:
  - There can be more than one protagonist per drama
  - Not only “heroes” in a positive sense, but also “anti-heroes” allowed
Experiments
<table>
<thead>
<tr>
<th>Feature Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>Character speech, normalised on whole text</td>
</tr>
</tbody>
</table>
## Features

<table>
<thead>
<tr>
<th>Feature Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>Character speech, normalised on whole text</td>
</tr>
<tr>
<td>Centrality</td>
<td>Different measures: <em>(weighted) degree, closeness, betweenness, eigenvector</em></td>
</tr>
</tbody>
</table>
## Features

<table>
<thead>
<tr>
<th>Feature Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>Character speech, normalised on whole text</td>
</tr>
<tr>
<td>Centrality</td>
<td>Different measures: <em>(weighted) degree, closeness, betweenness, eigenvector</em></td>
</tr>
<tr>
<td>Topic Modelling</td>
<td>10 Topics trained on the dramas</td>
</tr>
</tbody>
</table>
## Features

<table>
<thead>
<tr>
<th>Feature Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>Character speech, normalised on whole text</td>
</tr>
<tr>
<td>Centrality</td>
<td>Different measures: <em>weighted</em> degree, <em>closeness</em>, <em>betweenness</em>, <em>eigenvector</em></td>
</tr>
<tr>
<td>Topic Modelling</td>
<td>10 Topics trained on the dramas</td>
</tr>
<tr>
<td>Actives</td>
<td>Number of scenes where figure is present</td>
</tr>
</tbody>
</table>
# Features

<table>
<thead>
<tr>
<th>Feature Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>Character speech, normalised on whole text</td>
</tr>
<tr>
<td>Centrality</td>
<td>Different measures: <em>(weighted) degree, closeness, betweenness, eigenvector</em></td>
</tr>
<tr>
<td>Topic Modelling</td>
<td>10 Topics trained on the dramas</td>
</tr>
<tr>
<td>Actives</td>
<td>Number of scenes where figure is present</td>
</tr>
<tr>
<td>Passives</td>
<td>Number of Scenes where figure is mentioned</td>
</tr>
</tbody>
</table>
## Features

<table>
<thead>
<tr>
<th>Feature Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>Character speech, normalised on whole text</td>
</tr>
<tr>
<td>Centrality</td>
<td>Different measures: <em>(weighted) degree, closeness, betweenness, eigenvector</em></td>
</tr>
<tr>
<td>Topic Modelling</td>
<td>10 Topics trained on the dramas</td>
</tr>
<tr>
<td>Actives</td>
<td>Number of scenes where figure is present</td>
</tr>
<tr>
<td>Passives</td>
<td>Number of Scenes where figure is mentioned</td>
</tr>
<tr>
<td>LastAct</td>
<td>Is figure present in the last act?</td>
</tr>
<tr>
<td>Feature Name</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Tokens</td>
<td>Character speech, normalised on whole text</td>
</tr>
<tr>
<td>Centrality</td>
<td>Different measures: <em>(weighted) degree, closeness, betweenness, eigenvector</em></td>
</tr>
<tr>
<td>Topic Modelling</td>
<td>10 Topics trained on the dramas</td>
</tr>
<tr>
<td>Actives</td>
<td>Number of scenes where figure is present</td>
</tr>
<tr>
<td>Passives</td>
<td>Number of Scenes where figure is mentioned</td>
</tr>
<tr>
<td>LastAct</td>
<td>Is figure present in the last act?</td>
</tr>
<tr>
<td>NumberFig</td>
<td>In respective drama</td>
</tr>
</tbody>
</table>
# Features

<table>
<thead>
<tr>
<th>Feature Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>Character speech, normalised on whole text</td>
</tr>
<tr>
<td>Centrality</td>
<td>Different measures: <em>(weighted) degree, closeness, betweenness, eigenvector</em></td>
</tr>
<tr>
<td>Topic Modelling</td>
<td>10 Topics trained on the dramas</td>
</tr>
<tr>
<td>Actives</td>
<td>Number of scenes where figure is present</td>
</tr>
<tr>
<td>Passives</td>
<td>Number of Scenes where figure is mentioned</td>
</tr>
<tr>
<td>LastAct</td>
<td>Is figure present in the last act?</td>
</tr>
<tr>
<td>NumberFig</td>
<td>In respective drama</td>
</tr>
<tr>
<td>Genre</td>
<td>e.g. <em>Weimar Classicism, Bourgeois Tragedy, Naturalism, etc.</em></td>
</tr>
</tbody>
</table>
Figure 1: A) Betweenness centrality, B) Closeness centrality, C) Eigenvector centrality, D) Degree centrality. Source: https://en.wikipedia.org/wiki/Centrality
Experimental Setup

- 114 dramas in our corpus
Experimental Setup

- 114 dramas in our corpus
- Three annotators
Experimental Setup

- 114 dramas in our corpus
- Three annotators
- Each data point represents a character in a play
Experimental Setup

- 114 dramas in our corpus
- Three annotators
- Each data point represents a character in a play
  - Arrays of feature values
  - Class: P (Title character) / C (Not title character)
Experimental Setup

- 114 dramas in our corpus
- Three annotators
- Each data point represents a character in a play
  - Arrays of feature values
  - Class: P (Title character) / C (Not title character)
  - Example Emilia: \{tokens=0.09, actives=0.16, ..., class=P\}
Experimental Setup

- 114 dramas in our corpus
- Three annotators
- Each data point represents a character in a play
  - Arrays of feature values
  - Class: P (Title character) / C (Not title character)
  - Example Emilia: \{tokens=0.09, actives=0.16, ..., class=P\}
- Random forest
Results

3
## Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority BL</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>0.86</td>
</tr>
<tr>
<td>Tokens BL</td>
<td>0.62</td>
<td>0.99</td>
<td>0.76</td>
<td>0.93</td>
</tr>
<tr>
<td>Random Forest</td>
<td><strong>0.72</strong></td>
<td><strong>1.00</strong></td>
<td><strong>0.83</strong></td>
<td><strong>0.95</strong></td>
</tr>
</tbody>
</table>

**Table 1:** Results for classification of protagonists plus baselines. Shown are the average values of the experiments for each annotator.
Feature Distribution

Going a step back...

Figure 2: Feature distribution for one annotator’s data set.
Figure 3: Relative Feature Importance for one model.
Character Analysis
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince,” ruler of the region
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince,” ruler of the region
- **Marinelli**: Chamberlain of the prince

Summary of plot
- Emilia is engaged to Count Appiani
- The prince wants Emilia for himself
- His chamberlain Marinelli assassinates Appiani
- Orsina was the prince’s mistress and plots to kill him
- Odoardo kills Emilia at her wish
- The prince blames Marinelli for her death
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince,” ruler of the region
- **Marinelli**: Chamberlain of the prince
- **Odoardo & Claudia**: Emilia’s parents
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince,” ruler of the region
- **Marinelli**: Chamberlain of the prince
- **Odoardo & Claudia**: Emilia’s parents
- **Appiani & Orsina**: Count/Countess

Summary of plot
Emilia is engaged to Count Appiani
The prince wants Emilia for himself
His chamberlain Marinelli assassinates Appiani
Orsina was the prince’s mistress and plots to kill him
Odoardo kills Emilia at her wish
The prince blames Marinelli for her death
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince,” ruler of the region
- **Marinelli**: Chamberlain of the prince
- **Odoardo & Claudia**: Emilia’s parents
- **Appiani & Orsina**: Count/Countess

Summary of plot

- Emilia is engaged to Count Appiani
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince,” ruler of the region
- **Marinelli**: Chamberlain of the prince
- **Odoardo & Claudia**: Emilia’s parents
- **Appiani & Orsina**: Count/Countess

Summary of plot

- Emilia is engaged to Count Appiani
- The prince wants Emilia for himself
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince,” ruler of the region
- **Marinelli**: Chamberlain of the prince
- **Odoardo & Claudia**: Emilia’s parents
- **Appiani & Orsina**: Count/Countess

Summary of plot

- Emilia is engaged to Count Appiani
- The prince wants Emilia for himself
- His chamberlain Marinelli assassinates Appiani
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing

Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince”, ruler of the region
- **Marinelli**: Chamberlain of the prince
- **Odoardo & Claudia**: Emilia’s parents
- **Appiani & Orsina**: Count/Countess

Summary of plot

- Emilia is engaged to Count Appiani
- The prince wants Emilia for himself
- His chamberlain Marinelli assassinates Appiani
- Orsina was the prince’s mistress and plots to kill him
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- **Emilia**: Young bourgeois woman
- **Der Prinz**: “The prince,” ruler of the region
- **Marinelli**: Chamberlain of the prince
- **Odoardo & Claudia**: Emilia’s parents
- **Appiani & Orsina**: Count/Countess

Summary of plot

- Emilia is engaged to Count Appiani
- The prince wants Emilia for himself
- His chamberlain Marinelli assassinates Appiani
- Orsina was the prince’s mistress and plots to kill him
- Odoardo kills Emilia at her wish
Analysis of Single Characters

Example: Emilia Galotti by Gotthold Ephraim Lessing
Bourgeois tragedy

- Emilia: Young bourgeois woman
- Der Prinz: “The prince,” ruler of the region
- Marinelli: Chamberlain of the prince
- Odoardo & Claudia: Emilia’s parents
- Appiani & Orsina: Count/Countess

Summary of plot

- Emilia is engaged to Count Appiani
- The prince wants Emilia for himself
- His chamberlain Marinelli assassimates Appiani
- Orsina was the prince’s mistress and plots to kill him
- Odoardo kills Emilia at her wish
- The prince blames Marinelli for her death
Figure 4: Top seven figures with the highest token number in *Emilia Galotti*.
Figure 5: Active and passive presence in *Emilia Galotti*. A figure is only passively present in a scene if they are not actively present.
What has the model learnt about these characters?

**Figure 6**: Shapley graph for single figures in *Emilia Galotti*. Brackets mean: (Actual class – Predicted class).
What has the model learnt about these characters?

Figure 7: Shapley graph for single figures in *Emilia Galotti*. Brackets mean: (Actual class – Predicted class).
Shapley Graphs

What has the model learnt about these characters?

**Figure 8:** Shapley graph for single figures in *Emilia Galotti*. Brackets mean: (Actual class – Predicted class).
Shapley Graphs

What has the model learnt about these characters?

Figure 9: Shapley graph for single figures in Emilia Galotti. Brackets mean: (Actual class – Predicted class).
Shapley Graphs

What has the model learnt about these characters?

**Figure 10**: Shapley graph for single figures in *Emilia Galotti*. Brackets mean: (Actual class – Predicted class).
Shapley Graphs

What has the model learnt about these characters?

Figure 11: Shapley graph for single figures in *Emilia Galotti*. Brackets mean: (Actual class − Predicted class).
Shapley Graphs

What has the model learnt about these characters?

Figure 12: Shapley graph for single figures in *Emilia Galotti*. Brackets mean: (Actual class – Predicted class).
Take-away

- High performance for protagonist classification
Take-away

- High performance for protagonist classification
- Tendency to produce False Positives
Take-away

• High performance for protagonist classification
• Tendency to produce False Positives
• Tokens feature is strong but not sufficient
Take-away

- High performance for protagonist classification
- Tendency to produce False Positives
- Tokens feature is strong but not sufficient
- Analysis of single characters yields interesting insides
Appendix
Results

Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority BL</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>0.97</td>
</tr>
<tr>
<td>Tokens BL</td>
<td>0.38</td>
<td>1.00</td>
<td>0.55</td>
<td>0.95</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.46</td>
<td>1.00</td>
<td>0.63</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Table 2: Results for classification of title characters plus baselines.
Results

Experiment 3

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1woTokens</td>
<td>0.82</td>
<td>0.98</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>A2woTokens</td>
<td>0.78</td>
<td>1.00</td>
<td>0.88</td>
<td>0.96</td>
</tr>
<tr>
<td>A3woTokens</td>
<td>0.51</td>
<td>1.00</td>
<td>0.67</td>
<td>0.93</td>
</tr>
<tr>
<td>TFwoTokens</td>
<td>0.37</td>
<td>1.00</td>
<td>0.54</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table 3: Results without using tokens feature.
Figure 13: Correlation for centrality features.
Figure 14: Feature distribution.
Feature Importance

Figure 15: Relative Feature Importance.
## Confusion Matrix

<table>
<thead>
<tr>
<th>Ref</th>
<th>C</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>878</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>32</td>
<td>171</td>
</tr>
</tbody>
</table>

**Table 4:** Confusion matrix for A1.

<table>
<thead>
<tr>
<th>Ref</th>
<th>C</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1196</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>100</td>
<td>106</td>
</tr>
</tbody>
</table>

**Table 6:** Confusion matrix for A3.

<table>
<thead>
<tr>
<th>Ref</th>
<th>C</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>883</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>45</td>
<td>176</td>
</tr>
</tbody>
</table>

**Table 5:** Confusion matrix for A2.

<table>
<thead>
<tr>
<th>Ref</th>
<th>C</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pred</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1456</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>57</td>
<td>49</td>
</tr>
</tbody>
</table>

**Table 7:** Confusion matrix for TF.
Annotation

- Three annotators with overlapping and unique dramas

<table>
<thead>
<tr>
<th>Annotator</th>
<th># Dramas</th>
<th># Protagonists (%)</th>
<th># Non-Protagonists (%)</th>
<th># Figures Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>34</td>
<td>171 (16)</td>
<td>910 (84)</td>
<td>1081</td>
</tr>
<tr>
<td>A2</td>
<td>37</td>
<td>176 (16)</td>
<td>928 (84)</td>
<td>1104</td>
</tr>
<tr>
<td>A3</td>
<td>36</td>
<td>106 (8)</td>
<td>1296 (92)</td>
<td>1402</td>
</tr>
</tbody>
</table>

Table 8: Distribution of annotations.
Annotation

- Three annotators with overlapping and unique dramas

<table>
<thead>
<tr>
<th>Annotator</th>
<th># Dramas</th>
<th># Protagonists (%)</th>
<th># Non-Protagonists (%)</th>
<th># Figures Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>34</td>
<td>171 (16)</td>
<td>910 (84)</td>
<td>1081</td>
</tr>
<tr>
<td>A2</td>
<td>37</td>
<td>176 (16)</td>
<td>928 (84)</td>
<td>1104</td>
</tr>
<tr>
<td>A3</td>
<td>36</td>
<td>106 (8)</td>
<td>1296 (92)</td>
<td>1402</td>
</tr>
</tbody>
</table>

Table 8: Distribution of annotations.

<table>
<thead>
<tr>
<th>Combination</th>
<th># Dramas</th>
<th>Cohen’s $\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1+A2</td>
<td>6</td>
<td>0.83</td>
</tr>
<tr>
<td>A1+A3</td>
<td>6</td>
<td>0.46</td>
</tr>
<tr>
<td>A2+A3</td>
<td>7</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Table 9: Cohen’s $\kappa$ for different combinations of annotations.
## Results

### Experiment 1

<table>
<thead>
<tr>
<th>Data</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Majority Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>0.84</td>
</tr>
<tr>
<td>A2</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>0.84</td>
</tr>
<tr>
<td>A3</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
<td>0.92</td>
</tr>
<tr>
<td><strong>Tokens Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>0.72</td>
<td>1.00</td>
<td>0.84</td>
<td>0.94</td>
</tr>
<tr>
<td>A2</td>
<td>0.70</td>
<td>0.99</td>
<td>0.82</td>
<td>0.93</td>
</tr>
<tr>
<td>A3</td>
<td>0.44</td>
<td>1.00</td>
<td>0.61</td>
<td>0.91</td>
</tr>
<tr>
<td><strong>Random Forest</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>0.84</td>
<td>1.00</td>
<td>0.91</td>
<td>0.97</td>
</tr>
<tr>
<td>A2</td>
<td>0.80</td>
<td>1.00</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>A3</td>
<td>0.51</td>
<td>1.00</td>
<td>0.68</td>
<td>0.93</td>
</tr>
</tbody>
</table>

*Table 10: Precision, Recall and F1 for classifying protagonists and accuracy.*