Conceptual Approach

― Discovery of semantic connections to languages from named entities in the article‖

ELEVATE Framework [1]
- Recursive exploration of \textit{yago} relations
- \textbf{Country centric:}
  - \texttt{isCitizenOf}, \texttt{diedIn}, \texttt{isLocatedIn}, \texttt{isLeaderOf}, \texttt{isPoliticianOf}, \texttt{wasBornIn}, \texttt{livesIn}
- \textbf{Organization centric:}
  - \texttt{owns}, \texttt{created}, \texttt{worksAt}
- Linking of entities to languages

\textbf{Breadth-first-search (BFS):}
Stopping the exploration after the discovery of first language for a named entity

\textbf{Depth-first-search (DFS):}
Revealing all languages associated with a named entity exhaustively

ELEVATE Pipeline

\textbf{Event Data Collection} \quad \textbf{Named Entity Disambiguation} \quad \textbf{Entity-level Analytics} \quad \textbf{Semantic Aggregation} \quad \textbf{Spread Prediction}

\textbf{Objective:}
Automatically Predict the Event Diffusion into Foreign Language Communities

\textbf{Task:}
Pick the best candidates from all the scored languages

\textbf{Adjusted Thresholding}
- Threshold(θ) = average spread in the ground truth
- k-fold cross-validation
- Risk of picking the irrelevant languages

\textbf{Multi-label Classification}
- Output labels as the languages in event spread
- Candidate language scores as feature vectors
- Classifiers decide the spread

\textbf{Spread Prediction}

Other Research Works

\textbf{ELEVATE-Live} [3]
- Online news article virality prediction to countries
- Extension of ELEVATE framework

Available at: \url{https://elevate.greyc.fr}

\textbf{Semantic Fingerprinting} [2]
- Fine-grained entity-level Web content classification
- Concise semantic representation of documents based on their entities

\textbf{Publications}


\textbf{Short Biography}

3\textsuperscript{rd} year PhD student University of Caen, France Advisor: Prof. Marc Spaniol

\textbf{Research Interests}
- Entity-level analytics
- Data aggregation via LOD
- Deep learning for noisy data analytics

\textbf{Education}
- Master degree in Maths and Computing from Indian Institute of Technology (IIT) Patna (India)
- Bachelor degree in computer science from Guru Jambeshwar University of Science and Technology Hisar (India)

\textbf{Experimental Results}

Table:

<table>
<thead>
<tr>
<th></th>
<th>5-days</th>
<th>10-days</th>
<th>20-days</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>0.671</td>
<td>0.725</td>
<td>0.768</td>
</tr>
<tr>
<td>English</td>
<td>0.673</td>
<td>0.728</td>
<td>0.77</td>
</tr>
<tr>
<td>French</td>
<td>0.675</td>
<td>0.73</td>
<td>0.772</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.675</td>
<td>0.73</td>
<td>0.772</td>
</tr>
<tr>
<td>Italian</td>
<td>0.678</td>
<td>0.732</td>
<td>0.775</td>
</tr>
</tbody>
</table>

Macro-average scores for the adjusted threshold based models (4PC: number of predictions)

Table:

<table>
<thead>
<tr>
<th></th>
<th>5-days</th>
<th>10-days</th>
<th>20-days</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>0.671</td>
<td>0.725</td>
<td>0.768</td>
</tr>
<tr>
<td>English</td>
<td>0.673</td>
<td>0.728</td>
<td>0.77</td>
</tr>
<tr>
<td>French</td>
<td>0.675</td>
<td>0.73</td>
<td>0.772</td>
</tr>
<tr>
<td>Spanish</td>
<td>0.675</td>
<td>0.73</td>
<td>0.772</td>
</tr>
<tr>
<td>Italian</td>
<td>0.678</td>
<td>0.732</td>
<td>0.775</td>
</tr>
</tbody>
</table>

Macro-average scores for the machine learning approach (4PC: number of predictions)