Using text prediction for facilitating input and improving readability of clinical text

Magnus Ahltorp, Maria Skeppstedt, Hercules Dalianis, Maria Kvist

Method

Text prediction on a word level predicts what the user intends to type, given the previous context. Statistics from existing texts are used for the prediction, usually from the same domain. In this study, we evaluated the possibility of predicting clinical text using other medical text resources.

Results

Saved keystrokes were evaluated for \( n = 1 \).

<table>
<thead>
<tr>
<th>Evaluation data</th>
<th>Number of letters</th>
<th>Saved keystrokes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper ceiling: Läkartidningen (600 sentences)</td>
<td>64 560</td>
<td>32%</td>
</tr>
<tr>
<td>Frequent domain words, ICU notes (100 words)</td>
<td>675</td>
<td>20%</td>
</tr>
<tr>
<td>Frequent bigrams, ICU notes (100 bigrams)</td>
<td>870</td>
<td>35%</td>
</tr>
<tr>
<td>Frequent trigrams, ICU notes (100 trigrams)</td>
<td>1 453</td>
<td>33%</td>
</tr>
<tr>
<td>Medical exam questions mimicking clinical text (240 sentences)</td>
<td>13 535</td>
<td>26%</td>
</tr>
</tbody>
</table>

Future work

- Evaluate on a large corpus of authentic clinical text.
- Use more advanced text prediction techniques.
- Incorporate syntactic and semantic categories or frequent words or n-grams from authentic clinical text.
- Adapt/develop evaluation techniques to a fast typing user.

However, there are obstacles:

- authentic clinical text is not always available for application development and often contains sensitive information
- a text prediction tool trained on unmodified clinical text would also suggest abbreviations, common misspellings, etc.

Building a database of word and n-gram frequencies

A medical journal (Läkartidningen)

SNOMED CT, MeSH, ICD-10, FASS

Applying frequencies for retrieving \( n \) predictions

The user has typed: **Patient with diab**

Search for: "patient with diab" in tri-grams

\( n \) hits

\( n \geq n \) hits

Search for: "with diab" in bi-grams

\( n \) hits

\( n + n \geq n \) hits

Search for: "diab" in words

\( n \) hits

\( n' + n'' + n'' \geq 1 \)

Correspondence to: mariask@svsu.se