Example exam

Multimedia Retrieval 2016-2017, Utrecht University

Date: October 2016

Disclaimer:

1. No answers are provided. During the last practical, answers can be obtained. However, the best way to obtain the answers is to study the material and practice. If you do so, you’ll greatly benefit from this at the exam.

2. You are expected to have basic mathematical skills, in particular of linear algebra (e.g., calculations with matrices), which are needed to understand the material discussed in the book and the lectures.

3. In answering the questions, the answers must be complete, correct, and concise. Please also note that the (sub)questions can consist of multiple related questions.

4. This example exam is possibly slightly shorter (max. 33%) than the actual exam. Please let me know how much time it takes you to make this exam.

5. The aim is to assess your knowledge and skills on most of the course’s topics. However, possibly not all will be assessed. For example, in case of this example exam, research methods are not covered.

6. All questions are roughly equally weighted.

7. With both this example exam and the actual exam, you are allowed to use a calculator. You are not allowed to use other electronic devices (e.g., a smartphone), books, or other information carriers.

8. In case of errors or other problems, please contact me!

success!! Egon
1. [General]

(a) Define multimedia retrieval.
(b) Describe three of multimedia retrieval’s fields of application.
(c) Give three levels of abstraction, which can be used to specify multimedia content.

2. [Distance]

(a) What (formal) conditions must be satisfied for a distance to be a metric?
(b) Calculate the histogram intersection between vectors \(<1, 2, 3, 4>\) and \(<5, 6, 7, 8>\).
(c) Show or proof whether or not histogram intersection is a metric and, if not, define a metric form.

3. [Performance evaluation]

The \(F_\beta\)-measure is defined as:

\[
(1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}}.
\]

(a) Define precision and recall.
(b) Explain what the \(F_\beta\)-measure indicates and explain how we can put more emphasis on precision than recall (and the other way around) in the \(F_\beta\)-measure?
(c) Define an alternative performance measure for multimedia retrieval systems and explain its pros and cons compared to the \(F_\beta\)-measure.

4. [Signals]

(a) Define sampling, give an example in the audio domain, and explain the impact it can have on a signal.
(b) Define quantization, give an example in the image/video domain, and explain the impact it can have on a signal.
(c) The handbook defines two normalizations:

\[
\frac{x(l) - \min(x)}{\max(x) - \min(x)} \quad \text{and} \quad \frac{x(l) - \mu(x)}{\sigma(x)}
\]

Explain the behavior of both functions, discuss their downside, and compare both by discussing the difference(s).
5. [Information Retrieval]

(a) What is the relationship between LSA (Latent Semantic Analysis) and SVD (Singular Value Decomposition)?

(b) Assume we have a corpus (i.e., a set) of 3 documents (i.e., \(d_1, d_2, \) and \(d_3\)) containing a total of 4 words (i.e., \(w_1, w_2, w_3, \) and \(w_4\)). For this corpus, the following word-by-document (or term-by-document) table has been computed, where the cells contain the word frequencies:

<table>
<thead>
<tr>
<th></th>
<th>(d_1)</th>
<th>(d_2)</th>
<th>(d_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(w_1)</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>(w_2)</td>
<td>4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>(w_3)</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>(w_4)</td>
<td>6</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

SVD is applied to the original word-by-document table above, yielding a decomposition \(U \circ S \circ V'\) with \(U, S,\) and \(V\) rounded off for the purpose of this exercise as:

\[
U = \begin{bmatrix}
0.2 & 0.3 & 0.6 \\
0.4 & -0.7 & 0.3 \\
0.2 & 0.6 & 0.7
\end{bmatrix}, \quad S = \begin{bmatrix}
9 & 0 & 0 \\
0 & 4 & 0 \\
0 & 0 & 2
\end{bmatrix}, \quad V = \begin{bmatrix}
0.8 & 0.5 & -0.3 \\
0.6 & -0.7 & 0.2 \\
0.1 & 0.3 & 1.0
\end{bmatrix}
\]

Assume the corpus deals almost exclusively with two topics. Is it then reasonable to assume that the three dimensional document space can be mapped into a two dimensional semantic space?

(c) Suppose a user searches for information in the corpus using the words ‘\(w_1w_2\)’ as query. Show where the concept expressed by these words would lie in an (appropriately chosen) semantic space, by calculating the coordinates of the query \(w_1w_2\) in that space. (If you completely simplify your outcome, which is optional, round it off to no more than one decimal). Can the coordinates in the semantic space have negative values?