E9 205 Machine Learning for Signal Processing

Introduction to Machine Learning of Sensory Signals

Instructor - Sriram Ganapathy (sriramg@iisc.ac.in)
Teaching Assistant - Prachi Singh (prachisingh@iisc.ac.in)

http://leap.ee.iisc.ac.in/sriram/teaching/MLSP_19/
Feature Extraction

- Using measured data to build desirable values.
- Attributes of the data that are informative and non-redundant.
- Resilience to noise/artifacts.
- Facilitating subsequent learning algorithm.
Feature Extraction

- Representation Problem

Cartesian Coordinates

Polar Coordinates
Feature Extraction

Scope for this course

I. Feature Extraction in Text.

II. Feature Extraction in Speech and Audio signals.

III. Processing of Images.
Text Processing
Text Modeling - Introduction to NLP

❖ Definitions
❖ Documents, Corpora, Tokens (Terms)
❖ Term Frequency (TF)
❖ Collection Frequency (CF)
❖ Document Frequency (DF)
❖ TF-IDF
❖ Bag of words model
Example [Manning and Schutze, 2006]

<table>
<thead>
<tr>
<th>Word</th>
<th>cf</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>try</td>
<td>10422</td>
<td>8760</td>
</tr>
<tr>
<td>insurance</td>
<td>10440</td>
<td>3997</td>
</tr>
</tbody>
</table>

**Figure 6.7** Collection frequency (cf) and document frequency (df) behave differently, as in this example from the Reuters collection.

<table>
<thead>
<tr>
<th>term</th>
<th>$df_t$</th>
<th>$idf_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>car</td>
<td>18,165</td>
<td>1.65</td>
</tr>
<tr>
<td>auto</td>
<td>6723</td>
<td>2.08</td>
</tr>
<tr>
<td>insurance</td>
<td>19,241</td>
<td>1.62</td>
</tr>
<tr>
<td>best</td>
<td>25,235</td>
<td>1.5</td>
</tr>
</tbody>
</table>

**Figure 6.8** Example of $idf$ values. Here we give the $idf$’s of terms with various frequencies in the Reuters collection of 806,791 documents.
Perplexity

Measuring the goodness of language modeling

\[ PP(W) = P(w_1w_2\ldots w_N)^{-\frac{1}{N}} \]
\[ = \sqrt[N]{\frac{1}{P(w_1w_2\ldots w_N)}} \]

On a Wall-street Journal Corpus

<table>
<thead>
<tr>
<th>Unigram</th>
<th>Bigram</th>
<th>Trigram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perplexity</td>
<td>962</td>
<td>170</td>
</tr>
</tbody>
</table>

Speech and Audio Processing
Speech and Audio

❖ Speech/Audio - 1D signals
  ❖ Generated by pressure variations producing regions of high pressure and low pressure.
  ❖ Travels through a medium of propagation (like air, water etc).
❖ Human sensory organ - eardrum.
  ❖ Converting pressure variations to electrical signals.
  ❖ Action mimicked by a microphone.
Sound waves in a computer

- Analog continuous signal from the microphone
- Discretized in time - sampling.
- Digitized in values - quantization.

http://mlsp.cs.cmu.edu/courses/fall2014/lectures/slides/Class1.Introduction.pdf
Why do we need time varying Fourier Transform

- When the signal properties change in time
- DFT will only capture the average spectral character
- Short-window analysis can indicate the change in spectrum.
Summary of STFT Properties

\[ X[k, n_0] \]

http://en.wikipedia.org/wiki/Window_function
Narrowband versus Wideband

- Short windows - poor frequency resolution - wideband spectrogram
- Long windows - poor time resolution - narrowband spectrogram
Spectrogram of Real Sounds

time-domain

successive short DFTs

individual t-f cells merge into continuous image

Dan Ellis, “STFT Tutorial”