

cases	doc_1		doc_2		decision	id																					
	<table><tr><td>authors</td><td>• Schomaker, Lambert</td></tr><tr><td>title</td><td>Caveats on Bayesian and hidden-Markov models (v2.8)</td></tr><tr><td>publication_date</td><td>2016-08-18 00:00:00</td></tr><tr><td>source</td><td>SupportedSources.CORE</td></tr><tr><td>journal</td><td></td></tr><tr><td>volume</td><td></td></tr><tr><td>doi</td><td>None</td></tr><tr><td>urls</td><td>• https://core.ac.uk/download/518748347.pdf</td></tr><tr><td>id</td><td>id-1587066492160013619</td></tr><tr><td>abstract</td><td>This paper describes a number of fundamental and practical problems in the application of hidden-Markov models and Bayes when applied to cursive-script recognition. Several problems, however, will have an effect in other application areas. The most fundamental problem is the propagation of error in the product of probabilities. This is a common and pervasive problem which deserves more attention. On the basis of Monte Carlo modeling, tables for the expected relative error are given. It seems that it is distributed according to a continuous Poisson distribution over log probabilities. A second essential problem is related to the appropriateness of the Markov assumption. Basic tests will reveal whether a problem requires modeling of the stochastics of seriality, at all. Examples are given of lexical encodings which cover 95-99% classification accuracy of a lexicon, with removed sequence information, for several European languages. Finally, a summary of results on a non- Bayes, non-Markov method in handwriting recognition are presented, with very acceptable results and minimal modeling or training requirements using nearest-mean classification.Comment: Difference of v2.8 with v2.7: a) Final empirical (simulation) table for word-trie experiment with epsilon-in-the-probabilities; b) Some small text changes; c) used ispel</td></tr><tr><td>versions</td><td></td></tr></table>		authors	• Schomaker, Lambert	title	Caveats on Bayesian and hidden-Markov models (v2.8)	publication_date	2016-08-18 00:00:00	source	SupportedSources.CORE	journal		volume		doi	None	urls	• https://core.ac.uk/download/518748347.pdf	id	id-1587066492160013619	abstract	This paper describes a number of fundamental and practical problems in the application of hidden-Markov models and Bayes when applied to cursive-script recognition. Several problems, however, will have an effect in other application areas. The most fundamental problem is the propagation of error in the product of probabilities. This is a common and pervasive problem which deserves more attention. On the basis of Monte Carlo modeling, tables for the expected relative error are given. It seems that it is distributed according to a continuous Poisson distribution over log probabilities. A second essential problem is related to the appropriateness of the Markov assumption. Basic tests will reveal whether a problem requires modeling of the stochastics of seriality, at all. Examples are given of lexical encodings which cover 95-99% classification accuracy of a lexicon, with removed sequence information, for several European languages. Finally, a summary of results on a non- Bayes, non-Markov method in handwriting recognition are presented, with very acceptable results and minimal modeling or training requirements using nearest-mean classification.Comment: Difference of v2.8 with v2.7: a) Final empirical (simulation) table for word-trie experiment with epsilon-in-the-probabilities; b) Some small text changes; c) used ispel	versions		DUPLICATES		101
			authors	• Schomaker, Lambert																							
			title	Caveats on Bayesian and hidden-Markov models (v2.8)																							
			publication_date	2016-08-18 00:00:00																							
			source	SupportedSources.CORE																							
			journal																								
			volume																								
			doi	None																							
			urls	• https://core.ac.uk/download/518748347.pdf																							
			id	id-1587066492160013619																							
			abstract	This paper describes a number of fundamental and practical problems in the application of hidden-Markov models and Bayes when applied to cursive-script recognition. Several problems, however, will have an effect in other application areas. The most fundamental problem is the propagation of error in the product of probabilities. This is a common and pervasive problem which deserves more attention. On the basis of Monte Carlo modeling, tables for the expected relative error are given. It seems that it is distributed according to a continuous Poisson distribution over log probabilities. A second essential problem is related to the appropriateness of the Markov assumption. Basic tests will reveal whether a problem requires modeling of the stochastics of seriality, at all. Examples are given of lexical encodings which cover 95-99% classification accuracy of a lexicon, with removed sequence information, for several European languages. Finally, a summary of results on a non- Bayes, non-Markov method in handwriting recognition are presented, with very acceptable results and minimal modeling or training requirements using nearest-mean classification.Comment: Difference of v2.8 with v2.7: a) Final empirical (simulation) table for word-trie experiment with epsilon-in-the-probabilities; b) Some small text changes; c) used ispel																							
			versions																								