

5 ACKNOWLEDGEMENTS

Description	#Words	#Connections	#Epochs	Performance	Training Performance	Test
HUMAN	?	$\approx 10^{14}$?	?	79.4%	
ADPT	2640	13250	1	76.7%	47.1%	
ADPT_50	648	3240	1	96.0%	72.1%	
T_WP	648	4290	1	99.5%	72.1%	
	5694	4315	1	99.3%	68.6%	
	3240	26		99.9%	77.9%	
	9			99.5%	79.1%	

(see text for details)

Classifier that improves in

The EOL mails were labelled according to this division. All administrative information and routing information were automatically extracted except for the subject and the sender of a message. The results are summarized in Table 1.

2 Training Procedures

Two underlying training procedures were used for the ACCSYS-network:

Adaptive Learning: This algorithm was originally proposed by Gorin [1987]. It adjusts the weights w_{nk} between the n -th input unit v_n and the k -th output unit c_k based on the mutual information $I(c_k; v_n)$.

Flexibility Through Incremental Learning Neural Networks for Text Categorization

P. Geutner, U. Bodenhausen and A. Waibel

Department of Computer Science
University of Karlsruhe
7500 Karlsruhe 1
Germany

School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
USA

ogeutner@ira.uka.de, uli@cs.cmu.edu, ahw@cs.cmu.edu

Abstract

We present an adaptive incremental learning algorithm that learns interactively to classify messages (here: emails) into categories without the need for lengthy batch training runs. The algorithm was evaluated on a large database of email messages that fall into five subjective categories. As control experiment best human categorization performance was established at 79.4% for this task. The best of all connectionist architectures presented here achieves near human performance. This architecture acquires its language model and dictionary adaptively, either online or offline. The learning algorithm combines an adaptive phase for learning weights during interaction and a tuning phase for refining weights on static data. Such systems can be deployed in environments where necessary such as