In machine learning , sequence labeling is a type of pattern recognition task that involves the algorithmic assignment of a categorical label to each member of a sequence of observed values. A common example of a sequence labeling task is part of speech tagging , which seeks to assign a part of speech to each word in an input sentence or document. Sequence labeling can be treated as a set of independent classification tasks, one per member of the sequence. However, accuracy is generally improved by making the optimal label for a given element dependent on the choices of nearby elements, using special algorithms to choose the globally best set of labels for the entire sequence at once.

As an example of why finding the globally best label sequence might produce better results than labeling one item at a time, consider the part-of-speech tagging task just described. Frequently, many words are members of multiple parts of speech, and the correct label of such a word can often be deduced from the correct label of the word to the immediate left or right. For example, the word "sets" can be either a noun or verb. In a phrase like "he sets the books down", the word "he" is unambiguously a pronoun, and "the" unambiguously a determiner , and using either of these labels, "sets" can be deduced to be a verb, since nouns very rarely follow pronouns and are less likely to precede determiners than verbs are. But in other cases, only one of the adjacent words is similarly helpful. In "he sets and then knocks over the table", only the word "he" to the left is helpful (cf. "...picks up the sets and then knocks over..."). Conversely, in "... and also sets the table" only the word "the" to the right is helpful (cf. "... and also sets of books were ..."). An algorithm that proceeds from left to right, labeling one word at a time, can only use the tags of left-adjacent words and might fail in the second example above; vice versa for an algorithm that proceeds from right to left.

Most sequence labeling algorithms are probabilistic in nature, relying on statistical inference to find the best sequence. The most common statistical models in use for sequence labeling make a Markov assumption, i.e. that the choice of label for a particular word is directly dependent only on the immediately adjacent labels; hence the set of labels forms a Markov chain . This leads naturally to the hidden Markov model (HMM), one of the most common statistical models used for sequence labeling. Other common models in use are the maximum entropy Markov model and conditional random field .

See also

Artificial intelligence

Bayesian networks (of which HMMs are an example)

Classification (machine learning)

Linear dynamical system , which applies to tasks where the "label" is actually a real number

Machine learning

Pattern recognition

Sequence mining

References

Further reading

Erdogan H., [1] . "Sequence labeling: generative and discriminative approaches, hidden Markov models, conditional random fields and structured SVMs," ICMLA 2010 tutorial, Bethesda, MD

(2010)

This machine learning -related article is a stub . You can help Wikipedia by expanding it .