

Re: Neural net – guessing literature kind

Source: <http://coding.derkeiler.com/Archive/General/comp.programming/2003-11/1391.html>

From: osmium (r124c4u102_at_comcast.net)

Date: 11/19/03

Date: Tue, 18 Nov 2003 23:30:23 GMT

Borys Musielak writes:

- > *I have to write an algorithm which guesses the literature kind*
- > *having 100 words of text.*
- > *Of course the natural way to design it is by using a neural net*
- > *and back propagation.*
- > *The question is – how to design the net and what function to use*
- > *to make it work well. And even more important task is: what does really*
- > *differentiate poems from fiction, or criminals in terms of words (i*
- > *think of using some syntax loops like verses with rhymes).*
- > *Enough would be to differ poetry from prose. Better but i dont know if*
- > *easy would be guessing prose kind: like love story, whodoneit, fantasy*
- > *etc.*
- > *I dont expect solutions, I rather need some ideas and maybe links to*
- > *similar projects.*

It seems to me that poetry and prose differ mostly in line length and punctuation. I was taught that Shakespeare was poetry but I never quite bought that. It seemed more like an assertion than anything else.

For a prose classification scheme, I would think of something a bit like the Minnesota Multi-Phasic Personality Inventory (MMPI), which is probably well documented. As I see it it is a correlation problem. Discard the prepositions and articles and focus on the nouns mostly. If the word "tomahawk" is used it is one indicator of a western.