Using Supervised Learning to Solve Text-Based CAPTCHAs

Turhan Kimbrough
Department of Computer Science
Towson University
Towson, Maryland
tkimbr1@students.towson.edu

***Abstract*—CAPTCHA is an acronym for the "Completely Automated Public Turing test to tell Computers and Humans Apart". It is a mechanism which is used to distinguish real human users from bots. CAPTCHAs come in a variety of forms, including the deciphering of obfuscated text, transcribing of audio messages, tracking mouse movement, and more. This research will focus on automating the process of deciphering text-based CAPTCHAs using machine learning techniques. Specifically, supervised learning is used to develop neural networks capable of 100% accuracy for certain datasets. The goal of this research is to demonstrate the weaknesses associated with text-based CAPTCHA mechanisms, especially with the prevalence of machine learning tools.**

***Keywords*—machine learning, neural networks, supervised training, CAPTCHA**

Fig. 1. Example of a figure caption.

References

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.

[2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] K. Elissa, "Title of paper if known," unpublished.

[5] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.