

Using Artificial Neural Networks to Create and Assign Subject Headings to New Publications: A Review of the Literature

S. Ren Milmine, B.A., B. Ed., MLIS candidate, University of Alberta

Connect with me on LinkedIn



Why Do Artificial Neural Networks Matter to Libraries?

The process for assigning subject headings to new publications is time-consuming and relies on human understanding, limited to only a few people making a judgement call on a given work. Artificial neural networks, with the capacity to mimic intuitive decision-making through pattern recognition, may offer an efficient alternative that saves time and eliminates or minimizes individual bias. This review examines existing literature to determine if the possibility exists for artificial neural networks (ANNs) to accurately and reliably assign subject headings to new publications with little to no human intervention through focusing on three areas:

- **Artificial Neural Networks**—construction, current uses, potential for future applications
- **Subject Headings**—history, construction, controlled vocabulary, problems with bias
- **Automation**—history, development and implementation, current uses

By examining the literature concerning these three areas, the scope of overlapping elements between the three becomes evident, which contributes to how worthwhile it is to further investigate using ANNs to automatically assign subject headings to new publications at the Library of Congress, which would in turn affect all libraries that use LCSH for their own catalogues.

Figure 1. A simple diagram of a single perceptron (Chandra 2018).

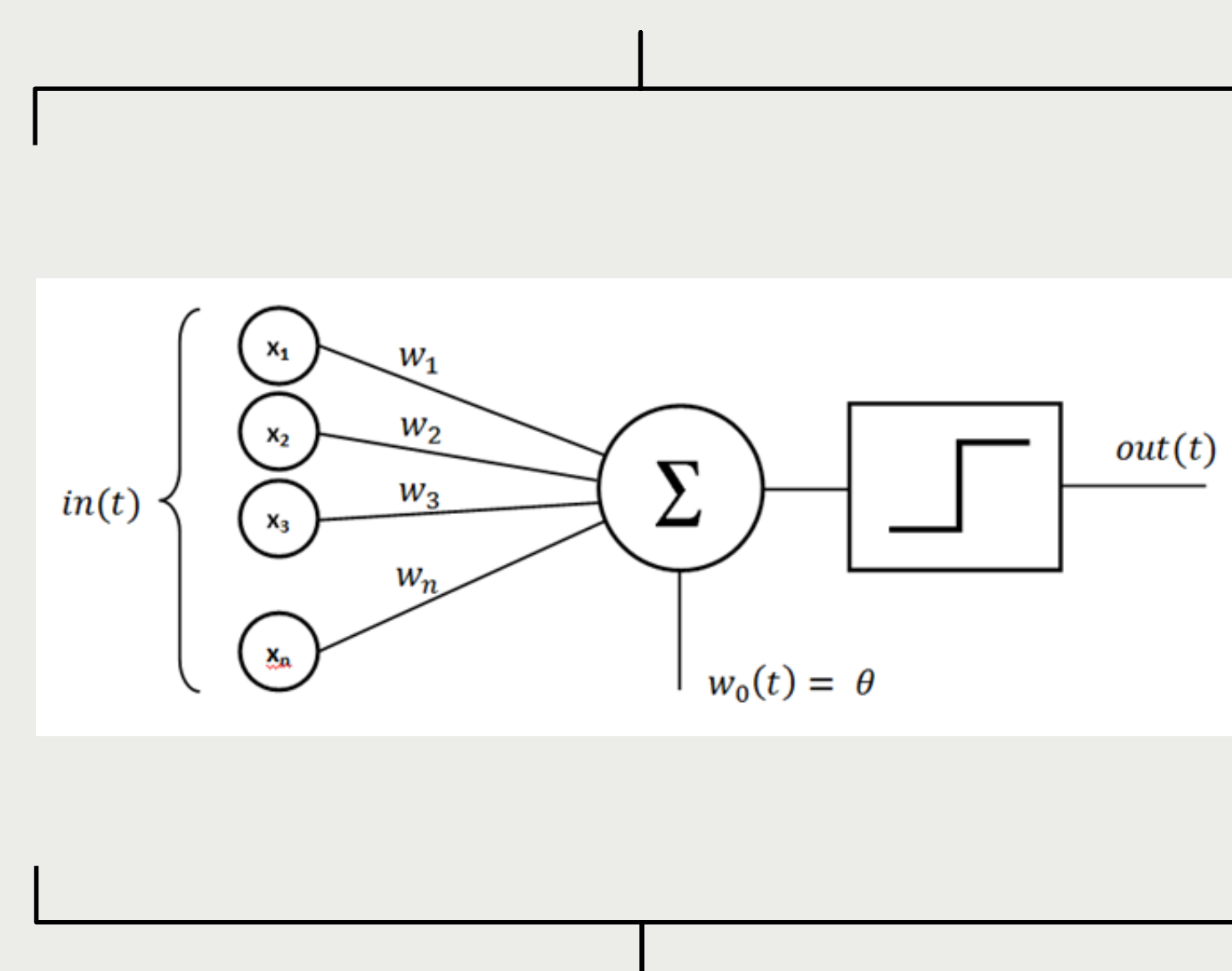
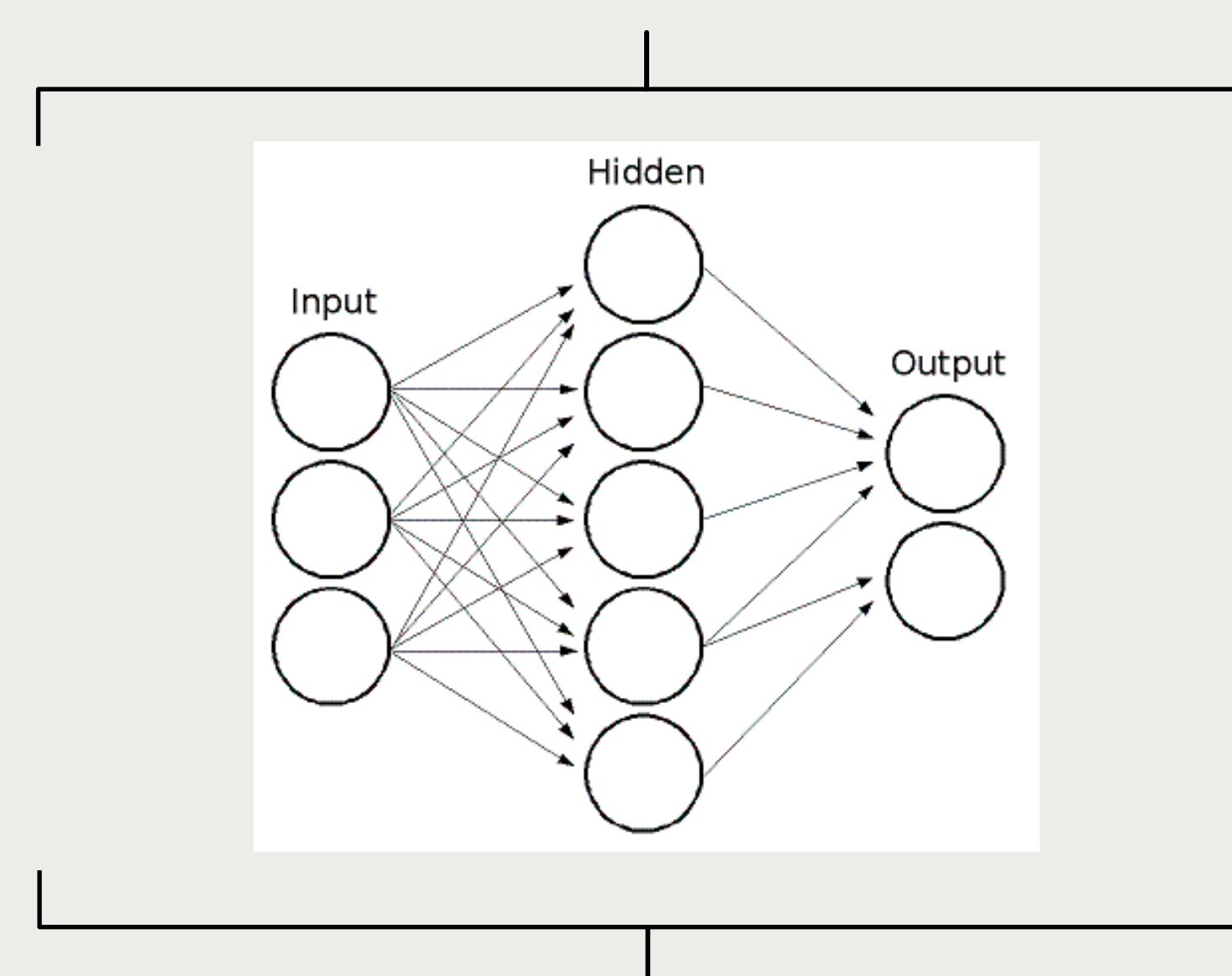


Figure 2. A simple diagram of a neural network (Shah 2017).



Selected References and Suggested Reading

- Beall, J. (2011). Academic Library Databases and the Problem of Word-Sense Ambiguity. *The Journal of Academic Librarianship*, 37(1), 64-69. doi:10.1016/j.acalib.2010.10.008
- Chandra, R. (2018, February 6). Neural networks for dummies: A comprehensive guide. *UpGrad BLOG*. Retrieved from <https://www.upgrad.com/blog/neural-networks-for-dummies-a-comprehensive-guide/>
- Golub, K. (2011). Automated Subject Classification of Textual Documents in the Context of Web-Based Hierarchical Browsing. *Knowledge Organization Vol 38*(3), p. 230-244.
- Goyal, P., Pandey, S., & Jain, K. (2018). *Deep Learning for Natural Language Processing: Creating Neural Networks with Python*. Berkeley, CA: Apress. doi:10.1007/978-1-4842-3685-7
- Lu, K., Mao, J., & Li, G. (2017). Toward effective automated weighted subject indexing: A comparison of different approaches in different environments. *Journal of the Association for Information Science and Technology*, 69(1), 121-133. doi:10.1002/asi.23912
- Nowick, E. A., Travnick, D., Eskridge, K., & Stein, S. (2010). A comparison of term clusters for tokenized words collected from controlled vocabularies, user keyword searches, and online documents. *Library Philosophy and Practice (e-journal)*. Retrieved from <http://digitalcommons.unl.edu/libphilprac/493>
- Shah, J. (2017, November 16) Neural networks for beginners: Popular types and applications: An introduction to neural networks learning. *Stats and Bots*, retrieved from <https://blog.statsbot.co/neural-networks-for-beginners-d99f2235efca>

The Literature Review: What Are the Results?

Overall, while there were no discussions specifically dedicated to using ANNs for automating subject heading creation or assignment, there was enough overlap between the three areas examined to justify further exploration into using ANNs for this type of work in libraries. Current developments occurring independently in each of these three areas like experiments with weighting and improvements in natural language processing serve to draw these three areas closer in purpose, function, and capability. What is now required are experts who are knowledgeable in both library-specific metadata needs and functional ANN construction methods to test practical applications of this potential solution to automating the creation and assignment of LC subject headings to new publications.

