

Redes neuronales artificiales desde un enfoque físico

Jesus David Prada Gonzalez
201214619

Director: Jaime E. Forero-Romero

November 4, 2016

1 Introduction

La inteligencia artificial es uno de los temas que han fascinado a la humanidad desde el desarrollo de la computación. Ésta ha sido tema recurrente de muchas de las historias modernas de ciencia ficción, y con los avances de la tecnología moderna en términos de poder computacional, esta idea se hace cada vez más palpable. Es por esto que la inteligencia artificial Multivac de Asimov es un personaje recurrente en sus grandiosos cuentos futuristas [1].

Uno de los aspectos de la inteligencia artificial que está al alcance de nuestras manos hoy en día, es el aprendizaje de máquinas, más conocido como Machine Learning. Éste hace referencia a la capacidad de las máquinas para aprender a enfrentarse a ciertos problemas sin ser específicamente programadas para ello. Ésto se logra por una variación de sus parámetros en lo que se conoce como aprendizaje supervisado (Supervised Learning), en donde se tiene una función de fitness o verosimilitud que indica qué tan acertado es el enfoque de la máquina al problema. Con varios intentos en el proceso de aprendizaje, la máquina puede calcular el gradiente aproximado de dicho fitness respecto a sus parámetros y modificarlos de tal manera que se optimice dicha función [2]. Así, si se entrena la máquina con un conjunto de datos conocidos, esta optimizará sus parámetros para poder enfrentarse a problemas nuevos de manera óptima.

Las dificultades de la programación de una inteligencia artificial se hacen claras al intentar programar, por ejemplo, un identificador de números en una imagen [3]. Lo que parece muy intuitivo para nosotros, como la identificación de curvas, contornos, intersecciones, etc. se vuelve un poco complicado al intentar expresarlo en un programa. De repente nos encontramos intentando clasificar casos, que en general son muy numerosos y harían el código muy ineficiente. De esta manera, deducimos que nuestra red neuronal, compuesta de diferentes capas de neuronas relacionadas entre sí, es óptima para enfrentarse a estos problemas, por esto uno de los abordajes más exitosos para Machine Learning reside en la mímica computacional de las redes neuronales biológicas.

De esta manera se pueden crear redes neuronales artificiales que procesen la información de imágenes para identificar objetos, rostros, animales, etc. Incluso, se pueden programar para que éstas aprendan a jugar videojuegos y más [4]. Fue de esta manera que se pudo crear una inteligencia artificial capaz de jugar Go y de vencer al campeón mundial.

A grandes rasgos, una neurona puede verse como un nodo que se activa o no dependiendo de unos estímulos de entrada. Computacionalmente esto se puede traducir a un output binario que se da como consecuencia de la aplicación de cierta función a unos inputs binarios. Los parámetros de dicha función serían los optimizados por el proceso de Supervised Learning. Teniendo esto en cuenta, una red de neuronas se forma al tomar como input de unas neuronas, el output de otras, formando así diferentes capas de procesamiento.

Lo curioso de estas redes neuronales es que son extremadamente eficientes al enfrentarse a ciertos problemas a pesar de que no se conoce muy bien el mecanismo de su éxito. Se sabe que ciertos diseños de redes funcionan de manera óptima para ciertos problemas, pero estas son conclusiones obtenidas heurísticamente. De esta manera, resulta importante poder cuantificar el proceso de aprendizaje de una neurona para monitorear su proceso de aprendizaje y poder obtener conclusiones de su mecanismo. Ya es conocido que muchas analogías de la mecánica estadística puede ser aplicadas al funcionamiento de las redes neuronales artificiales [?] y de esta manera se puede entender a grandes rasgos por qué funcionan al enfrentarse a problemas tan generales. De la misma manera, hay modelos estadísticos sobre redes neuronales biológicas [5], que pueden ser usados análogamente para modelar redes neuronales artificiales.

Consecuentemente, el objetivo de este proyecto de grado está en el estudio de modelos teóricos sobre redes neuronales artificiales para poder obtener conclusiones sobre el mecanismo de su éxito y verificarlo y aplicarlo a redes neuronales artificiales enfrentándose a problemas reales.

2 Objetivos Generales

Estudiar el comportamiento de las redes neuronales desde un enfoque físico para verificar si el mecanismo de su aprendizaje es verificable por medio de modelos teóricos.

3 Objetivos Específicos

- Desarrollar redes neuronales simples para familiarizarse con su diseño
- Estudiar a fondo las técnicas computacionales de diseño de redes neuronales, así

como los problemas técnicos a los que se pueden enfrentar

- Elegir entre las diferentes aplicaciones de redes neuronales una que sea versátil para el estudio de los modelos físicos asociados
- Estudiar qué posibles modelos físicos y estadísticos pueden ser aplicados a la red neuronal elegida para poder obtener conclusiones acerca de su funcionamiento
- Verificar las aplicaciones de estos modelos sobre la red elegida.
- Verificar la generalización de estos modelos para redes neuronales con diseños diferentes

4 Metodología

El estudiante realizará la investigación y el desarrollo de los temas individualmente, con el apoyo periódico del director por medio de las reuniones del grupo de Astrofísica. En estas obtendrá retroalimentación del desarrollo de su trabajo y se decidirá si en algún momento requiere de mayor asesoría por parte del director.

La metodología propuesta tiene un componente computacional fuerte relacionado con el desarrollo de varias redes neuronales y el monitoreo de su proceso de aprendizaje para la verificación de las conclusiones teóricas a obtener. A lo sumo, para agilizar los procesos de cómputo, se necesitará acceso al Cluster. Para el desarrollo de la parte teórica será necesaria la revisión de bibliografía especializada sobre el tema.

5 Cronograma

- Tarea 1: Desarrollo de redes neuronales simples
- Tarea 2: Investigación del diseño y funcionamiento de redes neuronales
- Tarea 3: Investigación de modelos teóricos sobre redes neuronales
- Tarea 4: Desarrollo de intuiciones o modelos verificables sobre el funcionamiento del aprendizaje de las redes
- Tarea 5: Preparación de la presentación de avance
- Tarea 6: Verificación de las conclusiones de los modelos anteriores sobre las redes neuronales artificiales específicas
- Tarea 7: Ampliación a redes neuronales más complejas

Tareas \ Semanas	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	X	X	X	X	X	X										
2	X	X	X	X	X	X	X	X	X							
3								X	X	X	X	X	X	X	X	X
4												X	X	X	X	X
5										X	X	X	X	X	X	X
9						X	X	X	X				X	X	X	X
Tareas \ Semanas	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
4	X	X	X	X	X	X										
6			X	X	X	X	X	X								
7						X	X	X	X	X	X					
8									X	X	X	X	X			
9	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

- Tarea 8: Desarrollo de resultados y conclusiones del trabajo
- Tarea 9: Escritura del documento final

6 Personas Conocedoras del Tema

- Juan Manuel Pedraza (Universidad de los Andes)
- Alonso Botero (Universidad de los Andes)
- Juan Gabriel Ramírez (Universidad de los Andes)

References

- [1] Asimov, I. (1974). The best of Isaac Asimov. Garden City, NY: Doubleday.
- [2] Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444. DOI: 10.1038/nature14539
- [3] M. (2016, January). Neural networks and deep learning. Retrieved November 04, 2016, from <http://neuralnetworksanddeeplearning.com/>
- [4] Lin, H. W., & Tegmark, M. (2016). Why does deep and cheap learning work so well? arXiv:1608.08225
- [5] Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. (2006, April 20). Weak pairwise correlations imply strongly correlated network states in a neural population. Nature, 440, 1007-1012.

Firma del Director

Firma del Estudiante