

Chat inteligente, la herramienta del futuro ¿Cuál es la mejor manera de crearla?

Smart chat, the tool of the future What is the best way to create it?

Nadine Castillo, Roymer Camacho, Luis Castilla, Katherine De la Hoz,
Universidad Simon Bolivar, Barranquilla

Resumen | Sabiendo que un chat inteligente es una poderosa herramienta informática que simula el comportamiento humano, siendo capaz de mantener una conversación coherente gracias al aprendizaje profundo de un determinado lenguaje; se realiza un análisis de los servicios que nos permiten crear este tipo de herramientas, se escogen las de mayor relevancia como son AMAZON, BLUEMIX y TENSORFLOW; llegando a la conclusión de que TENSORFLOW es la herramienta más accesible y fácil de usar debido a que puede empezar a utilizarse sin generar cobro.

Palabras clave: | Inteligencia Artificial, Maquinas de aprendizaje, Redes neuronales, Amazon, Bluemix, Tensorflow.

Abstract | Knowing that an intelligent chat is a powerful computer tool that simulates human behavior, being able to maintain a coherent conversation thanks to the deep learning of a certain language; An analysis of the services that allow us to create this type of tools is done, the most relevant ones are chosen such as AMAZON, BLUEMIX and TENSORFLOW; Concluding that TENSORFLOW is the most accessible and easy to use tool because it can start to be used.

Key words: | Artificial Intelligence, Machine Learning, Neural network. ks, Amazon, Bluemix, Tensorflow.

I. Introducción

El aprendizaje profundo es una forma de aprendizaje automático que se orienta hacia el avance de la inteligencia computacional. En el aprendizaje de máquina, se expone un equipo poco a poco a los nuevos datos durante un período de tiempo y se enseña a hacer predicciones basadas en esos datos. Luego, los desarrolladores se remontan en el software y hacen ajustes a los parámetros con el fin de mejorar la calidad de la predicción.

Cuando se trata de este tema novedoso aparecen diferentes servicios que facilitan a desarrolladores de todos los niveles de habilidad el uso de la tecnología de aprendizaje automático. En el siguiente artículo realizaremos el recuento del desarrollo de las actividades realizadas para la escogencia de la mejor herramienta entre:

AMAZON: Amazon Machine Learning – Amazon Lex
BLUEMIX: IBM WATSON
TENSORFLOW

Google emplea Deep Learning en sus algoritmos de reconocimiento de voz y reconocimiento de imagen basados en el software libre TensorFlow

En Amazon son pioneros aplicando estas técnicas en tareas como motores de recomendación, búsqueda de productos y detección de fraudes de comercio electrónico. Además están buscando científicos especializados en machine learning para expandirse a muchas otras áreas. Lo utilizan para adelantarse a tus gustos y los investigadores del MIT (Instituto Tecnológico de Massachusetts) para predecir el futuro.

Aprendizaje automático Amazon y aprendizaje automatico

Amazon Machine Learning es un servicio que facilita a desarrolladores de todos los niveles de habilidad el uso de la tecnología de aprendizaje automático. Amazon Machine Learning proporciona asistentes y herramientas de visualización que le guían a lo largo del proceso de creación de modelos de aprendizaje automático (ML) sin tener que aprender complicados algoritmos y tecnología de ML. Una vez que los modelos están listos, Amazon Machine Learning le permite obtener predicciones de su aplicación con facilidad utilizando API sencillas, sin tener que implementar código de generación de predicciones personalizado ni administrar infraestructuras. [1]

Amazon Machine Learning se basa en la tecnología ML probada y altamente escalable que la comunidad interna de científicos de datos de Amazon ha usado durante años. El servicio utiliza algoritmos potentes para crear modelos de ML al detectar patrones en los datos existentes. A continuación, Amazon Machine Learning utiliza estos modelos para procesar datos nuevos y generar predicciones para su aplicación. [2]

Amazon Machine Learning es altamente escalable y puede generar miles de millones de predicciones al día, además de abastecer dichas predicciones en tiempo real y con un desempeño excelente. Con Amazon Machine Learning no se necesita ninguna inversión inicial en hardware o software. Pagará lo que utilice, de modo que podrá comenzar con pocos recursos y escalar el servicio a medida que crezca su aplicación. [3]

IBM Watson Y Aprendizaje Automático

Watson es un sistema informático de inteligencia artificial que es capaz de responder a preguntas formuladas en lenguaje natural, desarrollado por la corporación estadounidense IBM. Forma parte del proyecto del equipo de investigación DeepQA, liderado por el investigador principal David Ferrucci. Lleva su nombre en honor del fundador y primer presidente de IBM, Thomas J. Watson.

Watson responde a las preguntas gracias a una base de datos almacenada localmente. La información contenida en esa base de datos proviene de multitud de fuentes, incluyendo enciclopedias, diccionarios, tesauros, artículos de noticias, y obras literarias, al igual que bases de datos externos, taxonomías, y ontologías [4].

Tensor flow y Aprendizaje Automático

Las aplicaciones de Google ahora son más inteligentes gracias al nuevo sistema de aprendizaje automático TensorFlow, que es más rápido, inteligente y flexible que su anterior sistema. Se puede adaptar con mayor facilidad a nuevos productos y es altamente escalable. Se puede ejecutar en un solo teléfono inteligente o en los miles de ordenadores en los centros de datos. Además, permite construir redes neuronales hasta cinco veces más rápidas que el anterior sistema de primera generación usado por Google.

El sistema de aprendizaje automático TensorFlow está especializado para el reconocimiento de voz, reconocimiento de textos y reconocimiento de imágenes, pero su inteligencia artificial todavía está en su infancia. Según Google todavía no puede hacer lo que un niño de 4 años hacer sin esfuerzo, como saber el nombre de un dinosaurio después de ver solamente un par de ejemplos, o entender que "vi el Gran Cañón volando a Chicago" no significa que el cañón estaba volando sobre la ciudad.

Google ha visto que TensorFlow tendrá un impacto aún más grande fuera de sus instalaciones y productos. Por ello, para que

su sistema de aprendizaje automático evolucione con mayor rapidez y hayan más aplicaciones inteligentes, han liberado el código fuente del proyecto TensorFlow para que investigadores y desarrolladores colaboren junto a Google para mejorar este gran sistema de inteligencia artificial. [5]

II. Desarrollo de la investigación

Los modelos basados en reglas facilitan que cualquier persona cree un chatbot. Pero es increíblemente difícil crear un bot que responda a consultas complejas. La coincidencia de patrones es un poco débil y por lo tanto, los bots basados en AIML (Artificial Intelligence Markup Language) sufren cuando encuentran una oración que no contiene patrones conocidos. Además, es mucho tiempo y requiere mucho esfuerzo para escribir las reglas manualmente. ¿Qué pasa si podemos construir un bot que aprende de conversaciones existentes (entre humanos)? Aquí es donde entra en juego el aprendizaje automático.

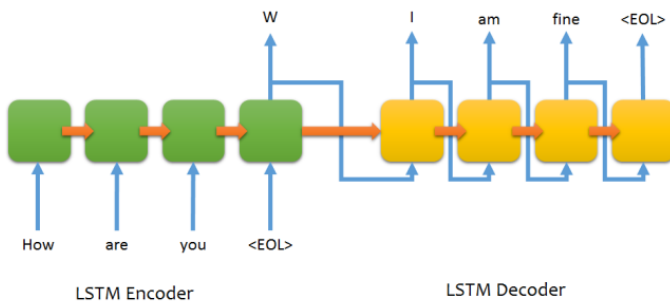
Llamemos a estos modelos que aprenden automáticamente de los datos, modelos inteligentes. Los modelos inteligentes pueden ser clasificados en:

Modelos basados en la recuperación
Modelos Generativos

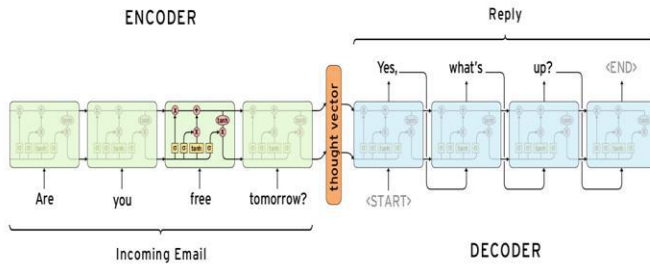
Los modelos basados en Recuperación seleccionan una respuesta de una colección de respuestas basada en la consulta. No genera nuevas oraciones, por lo tanto no necesitamos preocuparnos por la gramática [6]. Los modelos Generativos son bastante inteligentes. Generan una respuesta, palabra por palabra basada en la consulta. Debido a esto, las respuestas generadas son propensas a errores gramaticales. Estos modelos son difíciles de entrenar, ya que necesitan aprender la estructura de oración adecuada por sí mismos. Sin embargo, una vez entrenados, los modelos generativos superan a los modelos basados en recuperación en términos de manejo de consultas no vistas anteriormente y crear una impresión de hablar con un humano (un niño pequeño puede ser) para el usuario .

SEQ TO SEQ

El modelo Sequence To Sequence introducido en el aprendizaje de representaciones de frases utilizando RNN Encoder-Decoder para Statistical Machine Translation se ha convertido desde entonces en el modelo Go-To para sistemas de diálogo y traducción automática. Consiste en dos RNNs (Recurrent Neural Network): un codificador y un decodificador [7]. El codificador toma una secuencia (oración) como entrada y procesa un símbolo (palabra) en cada timestep. Su objetivo es convertir una secuencia de símbolos en un vector de características de tamaño fijo que codifica sólo la información importante en la secuencia mientras se pierde la información innecesaria [8-10].



Cada estado oculto influye en el siguiente estado oculto y el estado oculto final puede ser visto como el resumen de la secuencia. Este estado se llama contexto o vector de pensamiento, ya que representa la intención de la secuencia. Desde el contexto, el decodificador genera otra secuencia, un símbolo (palabra) a la vez.



III. Conclusiones

Luego de analizar de forma detallada las características de los servicios AMAZON MACHINE LEARNING, BLUEMIX y TENSORFLOW se encuentra que las tres herramientas tienen grandes funcionalidades, sin embargo, están dirigidas a distintos tipos de clientes ya que dos de ellas AMAZON y BLUEMIX generan cobro inmediato por su uso, mientras que TENSORFLOW escogida como la mejor herramienta, puede trabajarse inicialmente sin cobro, todo está en la necesidad y posibilidades de cada usuario. Adicionalmente fue posible determinar que para la creación de un chat inteligente es necesario utilizar el modelo generativo debido a que este es capaz de generar una respuesta, palabra por palabra basada en una consulta utilizando el modelo Sequence To Sequence que tiene como objetivo convertir una secuencia de símbolos en un vector de características de tamaño fijo que codifica sólo la información importante en la secuencia mientras se pierde la información innecesaria.

IV. Referencias bibliograficas

[1] J. Parra, L. Trujillo, and P. Melin, "Backpropagation learning with a (1+1) ES," presented at the Proceedings of the 12th annual conference companion on Genetic and evolutionary computation, Portland, Oregon, USA, 2010.
 [2] W. Choi *et al.*, "Hybrid network-on-chip architectures for accelerating deep learning kernels on heterogeneous manycore platforms," presented at the Proceedings of the

International Conference on Compilers, Architectures and Synthesis for Embedded Systems, Pittsburgh, Pennsylvania, 2016.

[3] J. Goncalves, V. Kostakos, and J. Venkatanathan, "Narrowcasting in social media: effects and perceptions," presented at the Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Niagara, Ontario, Canada, 2013.

[4] Z. Zhang, F. Masegla, R. Jain, and A. D. Bimbo, "KDD/MDM 2006: The 7th SIGKDD Multimedia Data Mining workshop report," *SIGKDD Explor. Newsl.*, vol. 8, no. 2, pp. 92-95, 2006.

[5] W. W. Zadrozny, S. Gallagher, W. Shalaby, and A. Avadhani, "Simulating IBM Watson in the Classroom," presented at the Proceedings of the 46th ACM Technical Symposium on Computer Science Education, Kansas City, Missouri, USA, 2015.

[6] G. Banavar, "Watson and the Era of Cognitive Computing," *SIGARCH Comput. Archit. News*, vol. 43, no. 1, pp. 413-413, 2015.

[7] O. Temam, "The rebirth of neural networks," *SIGARCH Comput. Archit. News*, vol. 38, no. 3, pp. 349-349, 2010.

[8] Y. Rivera Julio, "Bases De Datos Geográficas Y Autocorrelación Espacial Para Identificar Patrones De Distribuciones Espaciales," *Revista Investigación e Innovación en Ingenierías*, vol. 2 (1), 2014.

[9] E. De La Hoz, L. Lopez, y L. Perez, "Modelo de gestión de relaciones con los clientes en empresas de consultoría," *Revista Investigación e Innovación en Ingenierías*, vol. 5 (2), 2017.

[10] F. Rólon, «Opciones de la banca comercial en productos de importación,» *Dictamen Libre*, vol. 12/13, pp. 71-75, 2013.