Purchasing Request Text Classification

using Neural Network-based Algorithm

IPO Model Conceptual Framework

INPUT



PRDescription

กระบะลากหินทรายขนาด 600 ลิตร (Material : SS41) SG110D-08 "Cisco Gigabit Desktop Switch" BEARING 6301-ZZ SKF RTD SENSOR PT100 กระดิกน้ำร้อน ยี่ห้อ Sharp KP-Y40P ขนาด Oil Seal ID.75-90-10 mm./MAT:NBR เสาอากาศหนวดกังวัดความสงรถโม่ใหญ่ เทปโฟมกาวสองหน้า 21 มม.x5 ม. สก๊อตช์ 110 U-Bolt SUS304 2.1/4" x 5/16" รัดท่อ 2" BELT.V.Classical.C265 PSU.550W.MODEL V6.DELUX Safety relay output expansion 3SK1211 ข้อลดเกลี่ยวนอก hvd. 3/8" เกลียวใน1/4" เกลียว NPT MOELLER EXTENSION SHAFT TYPE NZM3/4-XV6(260193) Smart Clamp,2",Mat.SUS304,MJH(T),OD:60.5 BEARING, DEEP GROOVE, 6209/C3(F) SEAL, MECH, -, P-202, SEAL TECH. Magnatic ยี่ห้อ Totaline รุ่น P2820313 หัวอัดจาระบี 1/4"x19G-SUS304 CABLE ACC, 3RX8000-DCC32-1AFO [SD] RELAY, SOCKET, P2CF-08 OMRON ฝาราง เคเบิ้ลแลดเดอร์ 200 mm. PR-DCS,ANALOG INPUT,VALMET,D201736L "กระบะไม้ 2 ทางดีห่าง 29""x35"""

PROCESS



Word Embedding

(W2V model using Gensim library)



GRU model

OUTPUT



ClassSim

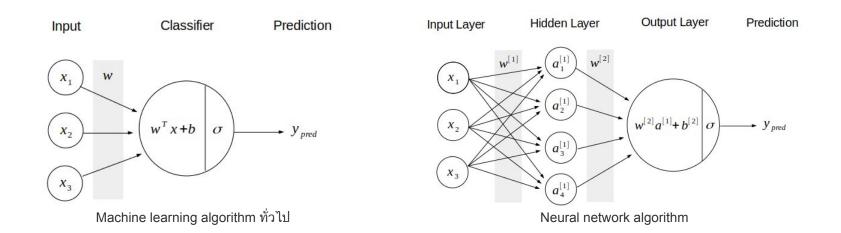
MTO
AV|Phone
Power Transmission
Oil Seal
Stationary
Power Transmission
IT Periferal
LV|Automation|Cable
Hydraulic
LV
Power Transmission

Air Condition
Hardware
LV|Automation|Cable
LV|Automation

Cable
DCS|Actuator
Pallet

Neural Network Algorithm

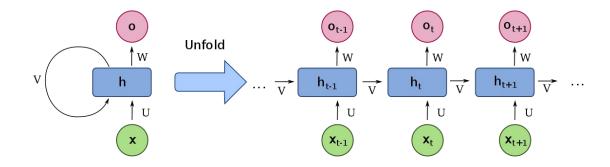
Artificial neural network เป็นวิธีการที่มีพื้นฐานมาจากการเลียนแบบการทำงานของสมองมนุษย์ ประกอบด้วยนิวรอน (Neurons) ที่ใช้ในการเชื่อมโยงเพื่อแก้ปัญหาต่าง ๆ ซึ่งเป็นการเรียนรู้แบบหนึ่งของ Machine learning



https://guopai.github.io/ml-blog14.html

Recurrent Neural Network (RNN)

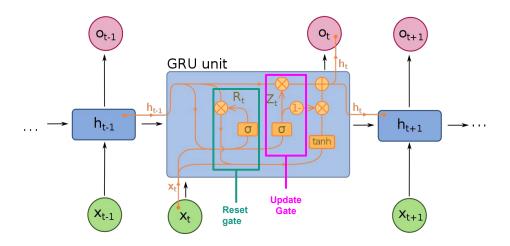
Recurrent Neural Network (RNN) คือ Artificial Neural Network แบบหนึ่งที่ออกแบบมาแก้ปัญหาสำหรับงานที่ ข้อมูลมีลำดับ Sequence โดยใช้หลักการ Feed สถานะภายในของโมเดล กลับมาเป็น Input ใหม่ คู่กับ Input ปกติ



A diagram for a one-unit recurrent neural network (RNN).

Gated Recurrent Unit (GRU)

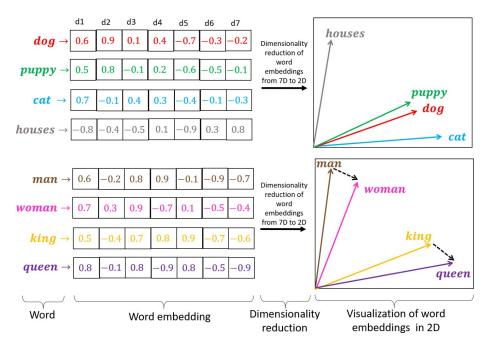
Gated Recurrent Units (GRU) เป็นกลไลปิดเปิดการอัพเดทสถานะภายใน Recurrent Neural Network



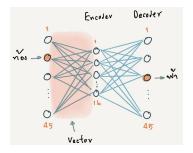
A diagram for a one-unit Gated Recurrent Unit (GRU).

Word Embedding

Word embeddings are vector representation of words learnt from context training.



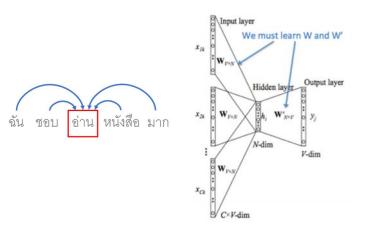
Word2Vec Model



การสร้าง Vector จากคำบริบทในประโยค โดยใช้เทคนิคทาง Neural Network แบบ Encoder-Decoder Model

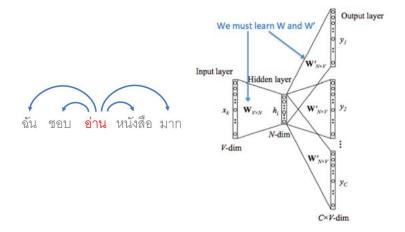
• Continuous Bag of Words (CBOW)

คำคำหนึ่งนั้นอาจสามารถถูกทำนายได้จากบริบทของคำที่อยู่รอบข้าง



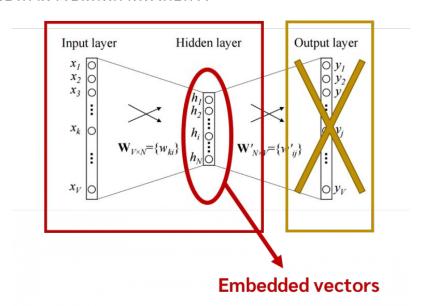
• Skip-gram

คำที่อยู่รอบข้างของ<mark>คำคำหนึ่ง</mark>อาจสามารถถูกทำนายได้จากคำคำนั้น



Word2Vec Model

สิ่งที่เราต้องการจากโมเดล Word2Vec คือ ค่าความสัมพันธ์ อธิบายได้ว่า การเอาผลลัพธ์จาก hidden layer หลังจากใส่ word vector ของคำมาใช้แทนคำนั้นโดยตรง



Word2Vec Model using Gensim Library

Gensim เป็น Open source Python Library ในกลุ่ม NLP

from gensim.models import Word2Vec

model = Word2Vec(sentences, vector_size = 100, window = 5, min_count = 5, workers = 3, sg = 1)

- vector_size (int, optional) Dimensionality of the word vectors.
- window (int, optional) Maximum distance between the current and predicted word within a sentence.
- min_count (int, optional) Ignores all words with total frequency lower than this.
- workers (int, optional) Use these many worker threads to train the model (=faster training with multicore machines).
- sg ({0, 1}, optional) Training algorithm: 1 for skip-gram; otherwise CBOW.

https://radimrehurek.com/gensim/models/word2vec.html

Python Scripts