Deep Learning Concepts and Applications Yoon Joong Kim, Dept. Of computer Engineering, Hanbat National University 2019.12.18

㈜넷비젼텔레콤

High-level Analogy with Astrophysics and coffee

High-Level Analogy with Astrophysics and Coffee



Replacing/Enhancing Iterative Algorithms with Neural Networks, Dadiv Wipf, Microsoft Research Beijing, ICASSP 2018 2

Similar for iterative algorithms and DNN?



Ultimate Dream : thinking machine



Ultimate dream : thinking machine



Schematic of a biological neu

Schematic of a biological neuron.

Neural model(Perceptron)



Logistic regression units





http://www.andreykurenkov.com/writing/a-brief-history-of-neural-nets-and-deep-learning/

Hardware implementations



Frank Rosenblatt, ~1957: Perceptron



Widrow and Hoff, ~1960: Adaline/Madaline

False Promises

• The Navy revealed the embryo of an electronic computer today that *it expects will be able to walk, talk, see, write, reproduce itself* and be conscious of its existence ... Dr. Frank Rosenblatt, a research psychologist at the Cornell Aeronautical Laboratory, Buffalo, said perceptrons might be fired to the planets as mechanical space explorers" July 08, 1958

The New York Eimes

http://query.nytimes.com/gst/abstract.html?res=9D01E4D8173DE53BBC4053DFB1668383649EDE

(Simple) AND/OR problem: linearly separable?

$$X \longrightarrow \overline{Y}$$



(Simple) AND/OR problem: linearly separable?



Perceptrons (1969) by Marvin Minsky, founder of the MIT AI Lab

- •In his book
- •We need to use MLP, multilayer perceptrons (multilayer neural nets)

$$X + b$$
 $\overline{W} + b$ $\overline{W} + b$ $\overline{W} + b$ $\overline{W} + b$ $\overline{Y} + b$

•No one on earth had found a viable way to train MLPs(W,b) good enough to learn such simple functions.

Frank Rosenblatt, ~1957: Perceptron Marvin Minsky, MLP (1969), unable to train w,b of MLP



"No one on earth had found a viable way to train"

*Marvin Minsky, 1969



http://cs231n.github.io/convolutional-networks/

Backpropagation (1974, 1982 by Paul Werbos, 1986 by Jeoffrey Hinton)



Jeoffrey Hinton, University of Toronto



Frank Rosenblatt, ~1957: Perceptron Marvin Minsky, MLP (1969), unable to train W,b of MLP Paul 1974/1982, Hinton 1986, Error Backpropagation

https://devblogs.nvidia.com/parallelforall/inference-next-step-gpu-accelerated-deep-learning/

Convolutional Neural Networks - Visual System by Hubel & Wiesel, 1959,62,58,...



Convolutional Neural Networks, [LeNet-5, LeCun 1980]





"At some point in the late 1990s, one of these systems was reading 10 to 20% of all the checks in the US."

[LeNet-5, LeCun 1980]

Fei-Fei Li & Andrej Karpathy & Justin Johnson Lecture 7 - 6 27 Jan 2016

Project NavLab 1984-1994, CMU



https://en.wikipedia.org/wiki/Navlab

Terminator 2 (1991)

JOHN: Can you learn? So you can be... you know. More human. Not such a dork all the time.



TERMINATOR: My CPU is a neural-net processor... a learning computer. But **Skynet** presets the switch to "read-only" when we are sent out alone.

We'll learn how to set the neural net

- **TERMINATOR** Basically. (starting the engine, backing out) The **Skynet** funding bill is passed. The system goes on-line August 4th, 1997. Human decisions are removed from strategic defense. **Skynet** begins to learn, at a geometric rate. It becomes **self-aware** at 2:14 a.m. eastern time, August 29. In a panic, they try to pull the plug.
- SARAH: And Skynet fights back.
- **TERMINATOR:** Yes. It launches its ICBMs against their targets in Russia.
- SARAH: Why attack Russia?
- **TERMINATOR:** Because **Skynet** knows the Russian counter-strike will remove its enemies here.

A BIG problem - number of layers-vanishing weights

- •Backpropagation just did not work well for normal neural nets with many layers
- •Other rising machine learning algorithms: SVM, RandomForest, etc.
- •1995 "Comparison of Learning Algorithms For Handwritten Digit Recognition" by LeCun et al. found that this new approach worked better



http://neuralnetworksanddeeplearning.com/chap6.html

CIFAR's contribution

- •Canadian Institute for Advanced Research(CIFAR)
- •CIFAR encourages basic research without direct application, was what motivated Hinton to move to Canada in 1987, and funded his work afterward.



"Everyone else was doing something different"

- "It was the *worst possible time*", says Bengio, a professor at universite de Montreal and co-director of the CIFAR program since it was rewarded last year. "Everyone else was doing something different. Somehow, Geoff convinced them."
- •"We should give(CIFAR) a lot of credit for making that gamble.
- CIFAR had huge impact in forming a community around deep learning





Hinton and Bengio's two breakthrough papers

- In 2006, Hinton, Simon Osindero, and Yee-Whye Teh published, "A fast learning algorithm for deep belief nets"
- Yoshua Bengio et al. in 2007 with "Greedy Layer-Wise Training of Deep Networks"

Breakthrough in 2006 and 2007 by Hinton and Bengio

- •In 2006, initializing weights Neural networks with many layers really could trained well, if <u>the weights are initialized in a clever way</u> rather than randomly.
- In 2007, DNN learning Deep learning methods are more efficient for difficult problems than shallow methods.
- •Rebranding to Deep Nets, Deep Learning

Large Scale Visual Recognition Challenge in IMAGENET



AlexNet of Doctoral research by Alex in Hinton's lab, 2010



System based on Deep learning, MSRA team 2015



Ensemble 2 by Trimps-Soushen(2016)

• Jie Shao, Xiaoteng Zhang, Zhengyan Ding, Yixin Zhao, Yanjun Chen, Jianying Zhou, Wenfei Wang, Lin Mei, Chuanping Hu

The Third Research Institute of the Ministry of Public Security, P.R. China.

- •Object classification/localization (CLS-LOC) Based on image classification models like Inception, Inception-Resnet, ResNet and Wide Residual Network (WRN), we predict the class labels of the image. Then we refer to the framework of "Faster R-CNN" to predict bounding boxes based on the labels. Results from multiple models are fused in different ways, using the model accuracy as weights.
- classification error : 2.99%

Neural networks that can explain photos



Deep API Learning

•Explain how to use API for a question



Figure 3: The Overall Workflow of DEEPAPI

copy a file and save it to -your destination path 35

FileInputStream.new FileOutputStream.new FileInputStream.getChannel File-OutputStream.getChannel FileChannel.size FileChannel.transferTo FileInput-Stream.close FileOutputStream.close FileChannel.close FileChannel.close

*GU et al. at HKUST with MSR/

Speech Recognition in noise environment



Game



Alphago Lee (2016)



Alphago Versions

Versions 🗢	Hardware 🗢	Elo rating +	Matches 🗢
AlphaGo Fan	176 GPUs, ^[52] distributed	3,144 ^[51]	5:0 against Fan Hui
AlphaGo Lee	48 TPUs, ^[52] distributed	3,739 ^[51]	4:1 against Lee Sedol
AlphaGo Master	4 TPUs, ^[52] single machine	4,858 ^[51]	60:0 against professional players; Future of Go Summit
AlphaGo Zero	4 TPUs, ^[52] single machine	5,185 ^[51]	100:0 against AlphaGo Lee 89:11 against AlphaGo Master
AlphaZero	4 TPUs, single machine	N/A	60:40 against AlphaGo Zero

Configuration and strength^[61]

프레임단위특징과 발음단위특징을 통합하는 Attention Mechanism을 이용한 음성감성인식 시스템

MSER(Merged Sepeech Emotion Recognition)



Automatic Bird-Species Recognition using the Deep Learning and Web Data Mining ,ICTC2018



Fig. 1. Flow chart of the Automatic Bird-Species Recognition.



Fig. 2. Left, header error due to data loss. Right, the white-background images.



Fig. 3. Left, outlier-removed images of birds. Right, outlier images of birds

GAN을 이용한 음성 감정 인식 모델의 성능 개선



Fig. 1. Proposed speech emotion recognition system architecture using GAN

그림 2. Attention과 BLSTM을 이용한 판별기 Fig. 2. Discriminator using attention and BLSTM

Real DataSet

Fake DataSet

DataSet

Two-Dimensional Attention-Based LSTM Model for Stock Index Prediction(ESCI, Scopus)



Fig. 1. Conceptual diagram of the 2D-ALSTM model composed of a CNN layer, an input and temporal attention layer, and an LSTM RNN layer.

A Voice Activity Detection Model composed of Bidirectional Long-Short Term Memory and Attention Mechanism.(IEEE Best Paper Award)



Fig.1. 10 seconds long noise added wave signal (SNR=10) and its ground truth for voice activity. Black indicates that the frame is labeled as voice.





http://www.wins.or.kr/papers/view.aspx?idx=189

ICTC 2018

- Bayesian Deep Learning-based Confidence-aware Solar Irradiance Forecasting System
 - HyunYong Lee and Byung Tak Lee (ETRI, Korea)
- Ensemble Classifier based on Decision-Fusion of Multiple Models for Speech Emotion Recognition
 - Kyoung-Ju Noh (ETRI, Korea)
- Distributed Deep Learning Framework based on Shared Memory for Fast Deep Neural Network Training
 - Eun-Ji Lim, Shinyoung Ahn, Wan Choi and Yoo-mi Park (ETRI, Korea)
- Automatic Bird-Species Recognition using the Deep Learning and Web Data Mining
- A development of a speech data transcription tool for building a spoken corpus
 - Hanbat National University
- Samples in ICTC 2018

Geoffrey Hitton's summary of findings up to today

- •Our labeled databases were thousands of times too small
- •Our computers were millions of times too small
- •We initialized the weights in a stupid way
- •We used the wrong-type of non-linearity

Why should I care?

•I am a researcher, not a computer scientist!

- •Do you have a idea?
- •Do you sell something?
- •Are doing any business?

Deep Learning Applications

•Yutube subtitle	*(자막)
E YouTube ^{AT}	terminator 2
	6
-	TensorFlow implementation
	<pre>copout_rate = tf.placeholder("float") _1 = tf.nn.relu(tf.add(tf.matmul(X, W1), B1)) _ = tf.nn.dropout(_L1, dropout_rate)</pre>
TRAIN: sess. dropo	<pre>run(optimizer, feed_dict={X: batch_xs, Y: batch_ys, ut_rate: (0.7}))</pre>
	TION: 시드라마, accuracy.eval({X: mnist.test.images, Y: dropout_rate: 1}) 자막
▶ ● 7:49 / 9:55	

•Facebook

f Search Facebook	۹ 🕯	Sung Home 👥 🛱 🙆 🖴	
Sung Kim Edit Profile FAVORITES News Feed Ads Manager	Vpdate Status Add Photos/Video Create Photo Album	YOUR ADS 누가 여법안을 발의 했나? 1 This Week	
Messages 15	Public - Post	Post Reach People Engaged	
 Photos 소프트웨어스토리 ISSTA 2 I'm a runner 누가 이법안을 받의 했 1 	ISSTA ISSTA Published by Andreas Zeller [?] - 45 mins - @ Doing a PhD in Software Testing and Analysis? Submit to the ISSTA 2016 Doctoral Symposium by April 22! Featuring a keynote by Alex Orso! Details: https://issta2016.cispa.saarland/doctoralsymposium/ Call for doctoral symposium submissions ISSTA is the leading research symposium in software testing and analysis, bringing together academics, industrial researchers, and cractifioners to exchange new	Recent Posts · 각 의원님들 보실때 이름 일의 bar로 표시된 처 Boost Post · · · · · · · · · · · · · · · · · · ·	
으, 시와 날씨 ♂ Top Photos 1 圓 TensorFlow KR			
Saved 11 PAGES FIEFE Transactions 20+	ideas, problems, and experiences on how to test and analyze software systems. The ISSTA Doctoral Symposium is a forum for PhD students work ISSTA2016.CISPA.SAARLAND		
Like Pages 14	38 people reached Boost Post	INVITE FRIENDS TO LIKE PAGES HKUST Water Sports Center Sports Center	
Create Page	🖒 Stevão Andrade 📰 👻	Amazon Web Services	

•Google search engine

Google (sung kim	♣ Q.	Serg III O 🚳
	All Images News Videos Maps More + Search tools		1 0 0
L	About 113,000,000 results (0.66 seconds) Sung Kim's CSE Homepage www.cse.ust.hk/=hunkim/ = Sung is an associate professor at the Hong Kong University of Science and Technology. He was a post-doc at the Program Analysis Group at MIT. He received Publications - Research - Software - Teaching Www.cse.ust.hk/=hunkim/Publications.html = Sung's Publications. 2015. Jaechang Nam and Sunghun Kim, "Heterogeneous Defect Prediction", In Proceedings of the 10th European Software Engineering Sung Kim - Wikipedia, the free encyclopedia Mtps://en.wikipedia.org/wiki/Sung_Kim = Sung Y. Kim (born 1990) is a Korean-born U.S. diplomat and the current United States Special Representative for North Korea Policy. He previously served as Early life and education - Professional career - Ambassador to South Korea		

•Amazon



• Display



- 5 Deep Learning Trends that will Rule 2019
 - Transfer learning
 - Transfer learning is widely popular machine learning technique, wherein a model, trained and developed for a particular task, is reused for performing another similar task.
 - VUI
 - A VUI (Voice User Interface or Vocal User Interface) is the interface for any speech application. Technically, VUI can either be a primary or a supplementary interface to enable natural voice communications between humans and machines.
 - ONNX(Open Neural Network Exchange) architecture
 - With ONNX, AI developers can more easily move models between state-of-the-art tools and choose the combination that is best for them." Proposed by Microsoft and Facebook in September-2017, ONNX models are now seeing increased adoption in commercial and open neural network libraries.
 - Machine comprehension
 - Machine Comprehension/Machine Reading Comprehension/Machine Reading are AI models that give a computer the ability to read a document and answer questions against it. While this is a relatively elementary task for a human, it is not that straightforward for AI models.
 - Edge intelligence
 - Edge Intelligence (EI) changes the way data is acquired, stored and extracted it shifts the process from the storage devices in the cloud to the edge (e.g. a camera, or a heat sensor)
 - AWS DeepLens "small Ubuntu- and Intel Atom-based computer with a built-in camera that is powerful enough to easily run and evaluate visual machine learning models."

- •15 Deep Learning Applications to Watch in 2019
 - Digital Assistants and Smart Devices
 - Machine Translation
 - Deep Learning Applications for Image Recognition
 - Self-driving cars
 - Digital Marketing
 - Customer Services
 - Better Recommendations
 - Robotics Deep Learning Applications
 - Colorizing Videos and Images
 - Deep Learning Applications Revolutionizing Medicine
 - Social Media
 - Marketing and Sales
 - Traveling
 - Healthcare
 - Smartphones

Thank you very much