

Paper Reading Roadmap

Neural Network

- General
  - Deep learning
  - Long Short-Term Memory
  - LSTM: A Search Space Odyssey
  - Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling
  - Sequence to sequence learning with neural networks
- Structure
  - Memory Networks
  - End-To-End Memory Networks
  - WaveNet: A Generative Model for Raw Audio
  - An Introduction to Variational Autoencoders
  - A Comprehensive Survey on Graph Neural Networks
- Learning Strategies
  - Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift
  - Dropout: A Simple Way to Prevent Neural Networks from Overfitting
  - ADAM: A Method for Stochastic Optimization
  - An overview of gradient descent optimization algorithms
  - Layer normalization
  - Group normalization

NLP

- General
  - Natural Language Processing (Almost) from Scratch
  - Advances in natural language processing
  - Recent trends in deep learning based natural language processing
- Topic Modeling
  - An introduction to latent semantic analysis
  - Probabilistic latent semantic analysis
  - Probabilistic topic models
  - Latent Dirichlet Allocation
- Representation Learning
  - A Neural Probabilistic Language Model
  - Distributed representations of words and phrases and their compositionality
  - Efficient Estimation of Word Representations in Vector Space
  - Glove: Global vectors for word representation
  - Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation
  - Enriching word vectors with subword information
  - Bert: Pre-training of deep bidirectional transformers for language understanding
  - Deep contextualized word representations
  - Improving language understanding by generative pre-training
  - Language models are unsupervised multitask learners
  - Language Models are Few-Shot Learners
- Classification
  - Convolutional neural networks for sentence classification
  - Deep learning for sentiment analysis: A survey
- Summarization
  - TextRank: Bringing Order into Texts
  - A Neural Attention Model for Abstractive Sentence Summarization
- Machine Translation
  - On the Properties of Neural Machine Translation: Encoder-Decoder Approaches
  - Effective Approaches to Attention-based Neural Machine Translation
  - Neural Machine Translation by Jointly Learning to Align and Translate
  - Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation
  - Attention is all you need
- Question Answering
  - VQA: Visual Question Answering
  - Ask Me Anything: Dynamic Memory Networks for Natural Language Processing
  - Squad: 100,000+ questions for machine comprehension of text
  - Know what you don't know: Unanswerable questions for SQuAD

Vision

- Classification
  - Imagenet classification with deep convolutional neural networks
  - Visualizing and understanding convolutional networks
  - Very Deep Convolutional Networks for Large-Scale Image Recognition
  - Going deeper with convolutions
  - Deep residual learning for image recognition
  - Densely Connected Convolutional Networks
- Object Detection
  - Overfeat: Integrated recognition, localization and detection using convolutional networks
  - Rich feature hierarchies for accurate object detection and semantic segmentation
  - Fast R-CNN
  - Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks
  - You Only Look Once: Unified, Real-Time Object Detection
  - YOLO9000: Better, Faster, Stronger
  - YOLOv3: An Incremental Improvement
  - YOLOv4: Optimal Speed and Accuracy of Object Detection
- Localization & Segmentation
  - U-Net: Convolutional Networks for Biomedical Image Segmentation
  - Learning deep features for discriminative localization
  - Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization

Machine Learning Basics

- The matrix calculus you need for deep learning
- Statistical Modeling: The Two Cultures
- Machine learning: Trends, perspectives, and prospects
- An introduction to ROC analysis
- Learning from imbalanced data
- Variational inference: A review for statisticians
- The expectation-maximization algorithm
- Dimension Reduction: A Guided Tour

Data Mining

- General
  - Interestingness Measures for Data Mining: A Survey
  - The PageRank citation ranking: Bringing order to the web
  - Process Mining Manifesto
- Pattern Mining
  - An Introduction to Variable and Feature Selection
  - Fast Algorithm for Mining Association Rules
  - A survey of sequential pattern mining
  - A Survey of Parallel Sequential Pattern Mining
- Clustering
  - A density-based algorithm for discovering clusters in large spatial databases with noise
  - Data Clustering: A Review
  - Techniques of Cluster Algorithms in Data Mining
  - Survey of Clustering Data Mining Techniques
  - On Clustering Validation Techniques
  - cValid: An R Package for Cluster Validation

Supervised Learning

- Kernel Machines
  - An Introduction to Kernel-based Learning Algorithms
  - A Tutorial on Support Vector Machine for Pattern Recognition
  - A Tutorial on Support Vector Regression
  - A Tutorial on nu-Support Vector Machines
- Ensemble
  - Bagging Predictors
  - Random Forests
  - A short introduction to boosting
  - Greedy Function Approximation: A Gradient Boosting Machine
  - Gradient Boosting Machine, A Tutorial
  - XGBoost: A Scalable Tree Boosting System
  - LightGBM: A Highly Efficient Gradient Boosting Decision Tree
  - CatBoost : unbiased boosting with categorical features
- Semi-supervised Learning
  - Combining Labeled and Unlabeled Data with Co-Training
  - Semi-supervised Learning with Deep Generative Models
  - Semi-Supervised Classification with Graph Convolutional Networks
  - MixMatch: A Holistic Approach to Semi-Supervised Learning
  - ReMixMatch: Semi-Supervised Learning with Distribution Alignment and Augmentation Anchoring
  - FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence

Unsupervised Learning

- Anomaly Detection: A Survey
- Deep Learning for Anomaly Detection: A Survey
- A Review of Novelty Detection
- LOF: Identifying Density-Based Local Outliers
- Support Vector Data Description
- Isolation Forest
- Isolation-based Anomaly Detection
- DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning

Artificial Intelligence

- General
  - Learning Deep Architectures for AI
  - Representation learning: A review and new perspectives
  - Generative Adversarial Networks
  - From evolutionary computation to the evolution of things
  - Probabilistic machine learning and artificial intelligence
  - AutoML: A Survey of the State-of-the-Art
- Reinforcement Learning
  - Human-level control through deep reinforcement
  - Mastering the game of Go with deep neural networks and tree search
  - An Introduction to Deep Reinforcement Learning
  - World Models
- Transfer Learning
  - Zero-shot learning through cross-modal transfer
  - Lifelong Learning with Dynamically Expandable Networks