

1. section1_強化学習

(1) エージェント

学習主体

(2) 環境

制御対象

(3) 状態

(4) 行動

報酬を最大化させるために、エージェントの行動を最適化する

(5) 報酬

エージェントの行動より得られる報酬

強化学習は静的なデータセットに依存せず、動的な環境で動作し。収集した経験から学習する。

データポイントすなわち経験は、環境とソフトウェアエージェントの間に試行錯誤のやり取りを通して学習中に収集される。

強化学習モデルは、監視なしにそれ自体で行動の学習を開始できる。

強化学習の複雑な問題は、ディープニューラルネットワークに依存している。→深層強化学習

自動運転では、カメラフレームや **Lider** 測定などの複数のセンサーを

同時に確認し、運転者に代わってニューラルネットワークがハンドルの切り方を決めることができる。

次のような産業用途に適している。

- ①高度な抑制
- ②自動運転
- ③ロボティックス
- ④スケジューリング
- ⑤キャリブレーション

2. section2_AlphaGo

Google DeepMind によって開発されたコンピュータ囲碁プログラムである。

- (1) AlphaGo Fan
- (2) AlphaGo Lee
- (3) AlphaGo Master
- (4) AlphaGo Zero
- (5) AlphaZero

人間が設定した評価経験則に従うのではなく、人間の棋譜を元に、コンピュータが自分自身との対戦を数千万回に渡り繰り返すことで強化していく。モンテカルロ木探索と呼ばれる探索アルゴリズムを組み合わせたことも特徴の一つである。

AlpaGo は盤面を評価するのに「value networks」、そして動きを選択するために「policy networks」という 2 つの人工知能に用いる評価関数をコンピュータ囲碁のアプローチとして使っている。

これらのディープニューラルネットワークは人間の囲碁のプロによる教師あり学習と、自己対戦による強化学習を行っておりその組み合わせにより囲碁の学習が進んでいく。

AlpaGo Zero はこれまでの人間の棋譜を元にするのではなく、囲碁のルールだけ教わり、その後は AI 同士の対局で独自の成長を遂げ、最強の打ち方を編み出した。

3. section3_軽量化_高速化技術

ディープラーニングを軽量化するモ「デル圧縮」 3 手法

(1) 高まるエッジ AI

現場のデバイス（エッジデバイス）に機械学習のモデルを実装し、その場で予測などを行うエッジ AI が注目を集めている。

- ・製造ラインの部品検査業務での異常検知らアクティビティセンシングなどリアルタイム応答が重要となるシーン
- ・自動運転で予想される通信不安定時の予測処理が可能になる。

エッジ AI を用いることで、エッジデバイスとクラウド間の通信が不要になるため、通信遅延や通信障害の影響を無くすることができる。通信コスト削減。

(2) エッジ AI の限界を払拭する「モデル圧縮」

エッジデバイスの処理能力の限界。

NVIDIA TensorRT や Amazon SageMaker Neo といったエッジ AI 向けの開発ツールらサービスが公開されている。

安価の評価用として NVIDIA Jetson Nano がある。

(3) モデル圧縮 3 手法

①Pruning（枝刈り）

ノード間の重みが小さい箇所の接続を削除する、または影響の小さいノードを削除する。

②Quantize（量子化）

重みなどのパラメータをより小さいビットで表現する。

TensorFlow や PyTorch などのディープラーニングのフレームワークでは、32ビットの浮動小数点制度を使用するのがほとんどであるが、8ビットの量子化であれば1%程度の性能低下であることが報告されている。

使用するビットを1~2ビットにすることでASICやFPGAなどでは計算効率が大きく向上する。

③Distillation（蒸留）

大きいモデルやアンサンブルモデルを教師モデルとして、その知識を小さいモデルの学習に利用する。これにより大きいモデルに匹敵する制度を持つ小さいモデルを作ることが期待できる。

4. section4_応用モデル

(1) MobileNet

① MobileNet V1

Depthwise & Pointwise Convolution

② MobileNet V2

Inverted residual block

③. MobileNet V3

Squeeze-and-Excite

(2) DenseNet

ResNet(ResBlock)と DenseNet(DenseBlock)の違い

ResBlock の場合、メイン側で畳み込みを行い、ショートカット側はあくまでバイパスを作って以前のレイヤーにおける値を足すだけである。

DenseBlock の場合、逆でショートカット側で畳み込みを行う。

(3) WaveNet

PixelCNN3 をベースとした音声波形を生成するための DNN の一つ。パラメトリック TTS モデルを採用。既存のパラメトリックモデルは通常、ボコーダとして知られる信号処理アルゴリズムに出力を渡すことで音声信号を生成する。

WaveNet は、オーディオ信号の生波形を一度に一サンプリングずつ直接モデル化することで、このパラダイムを変更している。

5. section5_Transformer

自然言語処理の世界では2018年に BERT が出てきた。

2019年には BERT を超える XLNet が出てきた。

- (1) 翻訳タスクにおいて、Seq2Seq よりも早くて精度が高い。
- (2) RNN、CNN も使わずに、Attention のみを使用した Encoder-Decoder モデルで、計算量も精度も改善。
並列処理も可能
- (3) アーキテクチャのポイント
 - ①エンコーダ・デコーダモデル
 - ②Attention
 - ③全結合層
- (4) NLP の最近の SoTA (BERT,XLNet,GPT-2 など) のベースモデル。

6. section6_物体検知_セグメンテーション

1) 画像認識

画像や動画データから特徴をつかみ、対象物を識別するパターン認識技術

2) 画像分類

機械学習やディープラーニングモデルで、画像を何らかの主題に基づき分類する処理方法

Xception

VGG

ResNet

Inception

MobileNet

DenseNet

3) 画像分類・物体位置特定

画像内のクラスと物体の位置を検出する方法

4) 物体検出

画像内のクラスと物体の位置を検出するが、複数種類の物体を1つの画像で検出すると、マルチクラスの物体検出

R-CNN

Fast R-CNN

SSD : Single Shot MultiBox Detector

Mask R-CNN

YOLOv3

RetinaNet

5) セグメンテーション

①. セマンティック・セグメンテーション

画像のピクセルを「どの物体クラス（カテゴリー）に属するか」で分類する。画像上の全ピクセルをクラスに分類する。

U-NET

MULTISCALE

HYBRID CNN-CRF

②. インスタント・セグメンテーション

画像のピクセルを「どの物体クラス（カテゴリー）に属するか、どのインスタンスに属するか」で分類する方法。物体毎の領域を分割し、かつ物体の種類を認識すること。

Mask R-CNN

DeepMask

FCIS

③. パノプティック・セグメンテーション

Semantic Segmentation と Instance Segmentation を組み合わせた方法。全てのピクセルにラベルが振られ、かつ数えられる物体に関しては、個別で認識した結果が返される。

Panoptic Feature Pyramid Network

UPNet