

グロッサリ

Glossary—グロッサリ—

隠れマルコフモデル

状態空間モデルの1つ。時系列信号を入力とし離散シンボル列の生成確率を出力する。マルコフモデルと異なり、状態系列が直接に観測されない（隠れている）。（篠田浩一）

n-gram 言語モデル

単語列において直前の（n-1）個の単語を与えたときのその次の単語の条件付確率を与えるモデル。nとしては3～5がしばしば用いられる。（篠田浩一）

強化学習

望ましい行動列を得たいときに、各々の行動に対して教師を与えるのではなく、各々の行動列にそれに対する報酬を与えて学習をする枠組みの総称。（篠田浩一）

DQN（Deep Q Network）

強化学習の方法の1つ。各行動に対する報酬の期待値を試行を繰り返して学習する枠組み（Q-learning）において、その期待値をニューラルネットワークの出力を用いて近似する。（篠田浩一）

表現学習

入力からその特徴表現を学習する枠組み。特に深層学習において、入力信号からそれに対応する固定長のベクトル（埋め込み表現）を得る手続きを指す。（篠田浩一）

分散表現

自然言語処理における意味の分散表現（distributed representation）は単語・句・文・段落等の意味を固定長ベクトル（実際には数百次元程度）で表現したものである。埋め込み（embedding）ともいう。大量テキストにおける文脈類似性に基づき、ニューラルネットワークを用いて分散表現を高速に計算する手法として、Word2Vecがよく知られている。（福島俊一）

Encoder-Decoder Model

入力系列→Encoder→分散表現→Decoder→出力系列，というステップで入力系列から出力系列への変換を行うモデルである。EncoderとDecoderにはニューラルネットワークを用いる。機械翻訳に使われたのをはじめ，自動要約，画像のキャプション生成等にも応用されている。（福島俊一）

ソフトクラスタリング

データをグループ化（クラスタリング）する際に，1つのデータが複数のグループに属してもよい条件で行うことをソフトクラスタリング，1つのデータは1つのグループにのみ属するという条件で行うことをハードクラスタリングという。（福島俊一）

F値

誤りの少なさを表す適合率（precision）と漏れの少なさを表す再現率（recall）はトレードオフ関係にある。適合率と再現率の調和平均はF値と呼ばれ，総合的なスコアとして用いられる。（福島俊一）

ミニバッチ

機械学習において，訓練データ集合からランダムに抜き出したサブセットで学習させる方法。訓練データ集合の規模が大きい場合，学習に大量の計算リソースを要するため，一部の訓練データを用いて近似するものである。（福島俊一）