

	単語	意味
day1	識別	データを目的のクラスに分類する
	生成	特定のクラスのデータを生成する
	ニューラルネットワーク	入力層と中間層、出力層からなる。重みとバイアスを最適化することで入力に対して出力の予想ができる。 回帰(結果予想やランキング)と分類(猫写真の判別、手書き文字の認識など)ができる。
	活性化関数	ニューラルネットワークにおいて、次の層への出力の大きさを決める非線形の関数。
	出力層	分類であれば各クラスごとの確率。
	誤差関数	二乗誤差の場合、教師データとニューラルネットワークの出力の誤差を二乗したもの。
	出力層における	値の強弱において、信号の大きさはそのままに変換。
	活性化関数	回帰では恒等写像、二値分類ではシグモイド関数、多クラス分類ではソフトマックス関数を使う。
	勾配降下法	ニューラルネットワークを学習させるための手法。前回の結果と間違えた分を加味していく。
	エポック	入力層、出力層を通じて出力層と訓練データの誤差を加味して重みとバイアスを変更するまでの一周のこと。
	確率的勾配降下法(SGD)	学習に使うデータをランダムに抽出したサンプルを使う。オンライン学習。
	勾配降下法	全サンプルの平均誤差。バッチ学習
	オンライン学習	モデルに都度都度データを追加してパラメータを更新していく。
	バッチ学習	全てのデータを一気に使ってパラメータ更新を行う。
	ミニバッチ勾配降下法	ランダムに分割したデータの集合(ミニバッチ)に属するサンプルの平均誤差。 小分けして一度にたくさんのバッチに対して同時並行でできることがメリット。SIMD。
	誤差逆伝播法	算出された誤差を、出力層側から順に微分し、前の層へと伝播していく。
	データ拡張	学習データが不足するときに人工的にデータを作り水増しする方法。分類タスク(画像認識)に効果が高い。 既存のデータを回転したりぼかしたりする。

day2	勾配消失問題	誤差逆伝播法が下位層に進んでいくにつれて勾配がどんどん緩やかになる。
	シグモイド関数の問題	シグモイド関数を微分すると最大値は0.25になる。つまり、中間層が増えるほど小さくなる。
	勾配消失問題の解決法	活性化関数の選択(ReLU関数:微分すると0か1)、重みの初期値設定、バッチ正規化
	バッチ正規化	ミニバッチ単位で、入力値のデータの偏りを抑制する手法。
	モメンタム	誤差をパラメータで微分したものと学習率の積を減産した後、現在の重みに前回の重みを減算した値と慣性の積を加算する。 局所的最適解にはならず、大域的最適解になる。谷間についてから最適値に行くまでの時間が早いというメリットがある。
	AdaGrad	誤差をパラメータで微分したものと再定義した学習率の積を減算する。
	RMSProp	誤差をパラメータで微分したものと再定義した学習率の積を減算する。大域的最適解となる。
	Adam	モメンタムとRMSPropを合体させたもの。
	正則化	ネットワークの自由度を下げて過学習を抑制する。
	CNN	次元間でつながりのあるデータを扱える。音声画像動画など。
	パディング	フィルターを使う前に、入力画像に縦横方向にピクセルを追加する。出力画像を入力画像と同じ画像サイズにできる。
	ストライド	フィルターを動かす際に、ずらし方を一個飛ばしのようにずらし方を変えることで出力の画像サイズを小さくする。
	チャンネル	フィルターの数を示す。
	プーリング層	対象領域の最大値または平均値を取得する。