

AI 自習ドリル

第15回 ディープ・ラーニングの1つ… 1つだけの結果が欲しいときに使う「回帰問題」

牧野 浩二

今回は回帰問題を扱います。ディープ・ラーニングと言うと分類問題をイメージされる方が多いと思いますが、この回帰問題も重要な問題設定です。分類問題は複数の出力ノードから1番値が大きいものを選択するといったものですが、出力ノードの数が多くなると扱いにくいといったデメリットがあります。これを解

決するのが回帰問題です。回帰問題として扱えば出力ノードが一意に定まるため、分類問題では複雑になりそうな問題でも解くことができます。本稿では、さまざまな回帰問題を設定し、ディープ・ラーニングで解いてみます。

1 できること

回帰問題は、回帰分析という手法を用いる問題で、これはたくさんのデータから数値を予測するものです。例題として次がよく用いられます。

- 身長と体重との関係
- 人口密度と薬局数との関係
- 駅の平均乗降客数とスーパーの売上高との関係
- 店舗面積 / 従業員数 / 商圈人口 / 広告費と来客数
- エンジンのスペックや車体重量、年式と燃費

● 単回帰分析

例えば身長と体重をもとに図1-1のように、たくさんのデータから直線を引いて、その直線を利用して体重から身長を予測することができます。

図1-1は直線で引けました。例えば駅の平均乗降

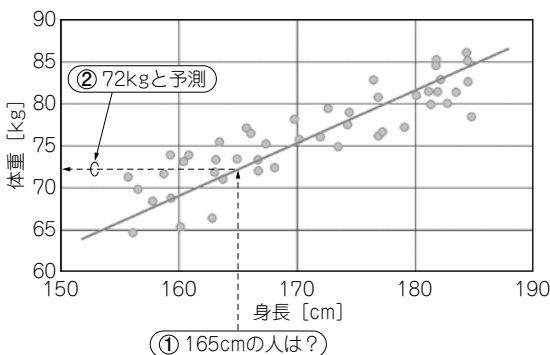


図1-1 身長から体重を予測する例

客数とスーパーの売上高の関係を考えます。乗降客数が少なすぎると売り上げは少ないですが、逆に多すぎても売り上げが少なくなるような傾向があることが図1-2から読み取れます。これに例えば2次関数を当てはめて、どの駅に出店するのがよいかを見ることができます。

このように2つの要素であれば、グラフにできます。これは単回帰分析と呼ばれます。

● 重回帰分析

図1-3のように店舗面積 / 従業員数 / 商圈人口 / 広告費といったさまざまな要因が重なった場合、グラフに表すことができないうので問題が難しくなります。これは重回帰分析と呼ばれています。

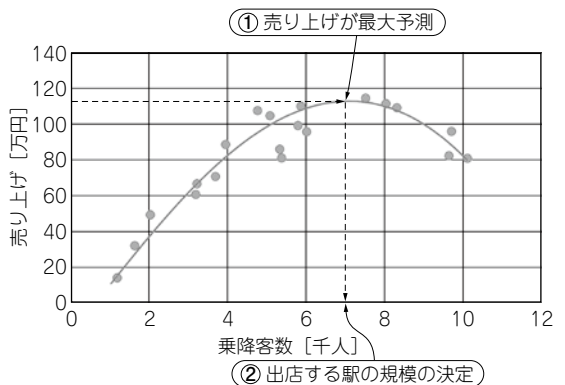


図1-2 売上げ最大を取れる人数から出店する駅を選ぶ