

機械学習のしくみ

GLEAP 機械学習講習会

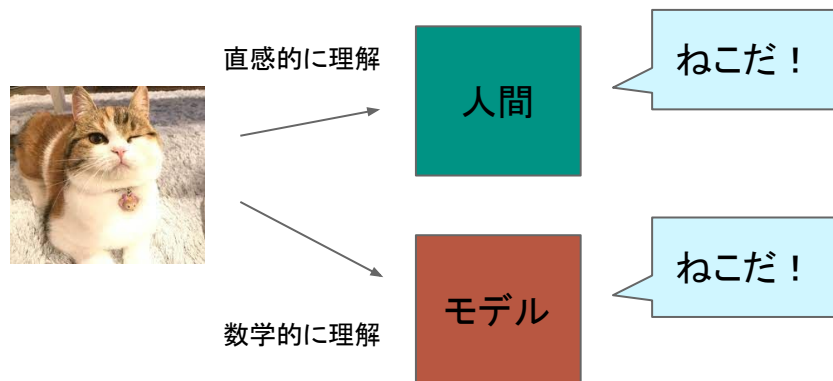
機械学習とは？

人工知能研究の1分野として「人間の学習能力や認知・判断能力を計算機上で再現する」こと

データに潜む構造・パターンを捉えること

より詳しく言うと、、、

データを**数学的に**解釈しモデルに適合させるための調整可能な**パラメータ**を決定すること



人間: 生まれて今までの間、何回もねこを見てきているのでこれが猫だとわかる

機械: なにもわからない

→ たくさんの猫の画像から数学的にパターンを学習 (パラメータ調整)

以上からなんとなく機械学習の流れがわかったかと思いますが...

要するに、

① 機械がデータを学習

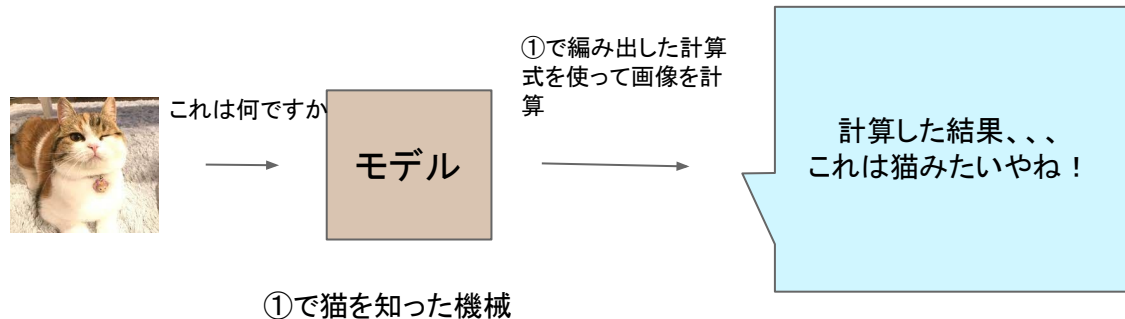


この画像をピクセルごとに分解してrgbの値を配列にしてみるとこうなっているのかー、
これに4と3と5と7をかけてこの関数をかけて、また2と4と6をかけて....(略)
よし、この計算式を使うと全部猫って判定できるみたいやで〜(早口)

モデルの頭の中はこんな感じです

当たり前ですが機械は**数字**としてしか物を認識できません。

② 機械が判定



世間でAIと呼ばれているものは、全部このようにただ計算が行われているだけです。

AIってなんか凄そうレベルでしかわかっていなかった人も少しは概念を理解できたかと思います。

機械学習の事前準備

では、機械学習モデルを作る準備として我々は何を準備したらいいのでしょうか？

① やりたいこと

例: 猫と犬の分類、株価予測

② 使用するモデル(アルゴリズム)

例: k近傍法, 線形回帰, ランダムフォレスト

研究が盛んで、現在無限とっていいほど様々なものがある。

③ データセット

例: 大量の犬と猫の画像, これまでの株価の時系列データ

④ その他

例: 時間、やる気、エラーに屈しない気合

機械学習の種類

- ・教師あり学習

ラベル付き学習データの基について予測するモデル

- ・分類

2つ以上の離散カテゴリ(犬か猫かなど)
の分類ラベル予測

- ・回帰

連続ラベルを予測するモデル
(値段などの連続的な値)

- ・教師なし学習

ラベルのないデータの構造を識別するモデル

- ・クラスタリング

データ内のグループを検出するモデル

※ラベルというのは正解データのこと
例えば、



ラベル: 猫

機械学習の種類

・教師あり学習

ラベル付き学習データの基づいて予測するモデル

・分類

2つ以上の離散カテゴリ(犬か猫かなど)
の分類ラベル予測

・回帰

連続ラベルを予測するモデル
(値段などの連続的な値)

・教師なし学習

ラベルのないデータの構造を識別するモデル

・クラスタリング

データ内のグループを検出するモデル

※ラベルというのは正解データのこと
例えば、



ラベル: 猫