第3回
y=ax+bから始める
初心者向けML講座

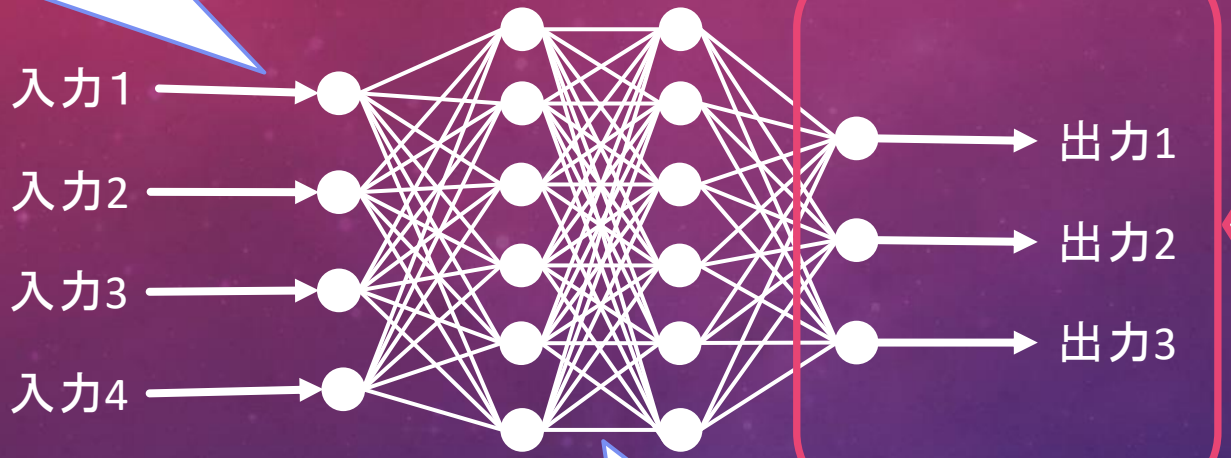
これまでのあらすじ

ディープラーニングの基本はニューラルネットワーク（NN）

ML Shukai

大量のデータを入れて
学習すると
正しい答えを予測してくれる

入力値と重みのパラメータ
を使って予測値を計算



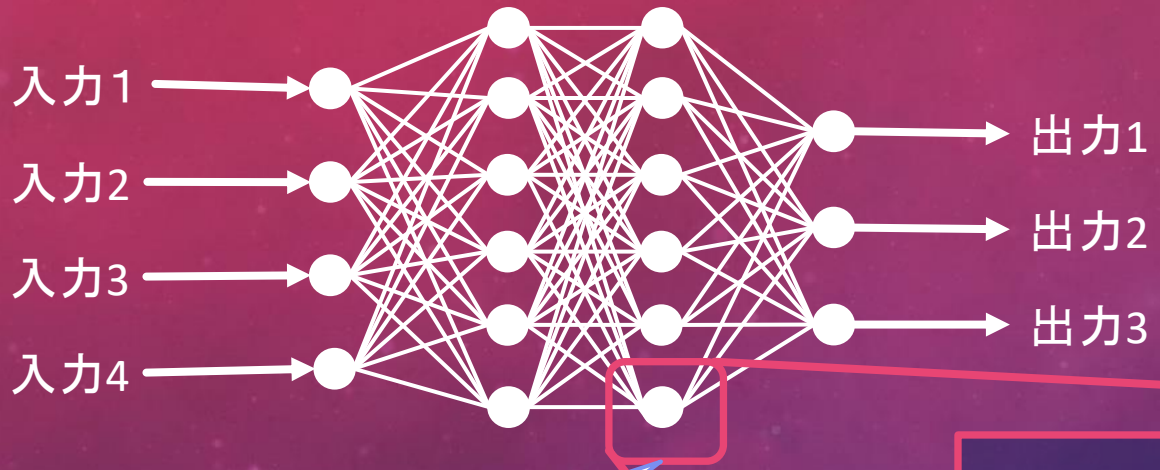
誤差

入力データ
とセットの
正解値

偏微分を行い、
正解値と予測値から
線1本1本の重みを更新
(学習)

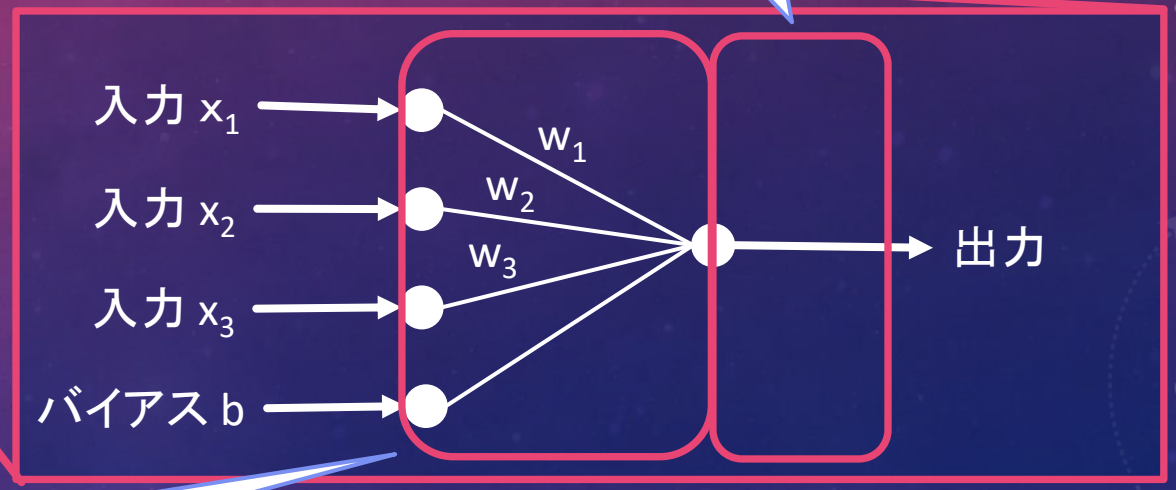
学習した結果
誤差が小さくなれば結果的
に正しい予測
をしてくれるAIになる

ディープラーニングの基本はニューラルネットワーク（NN）



ノード＝パーセプトロン

活性化関数を通して出力



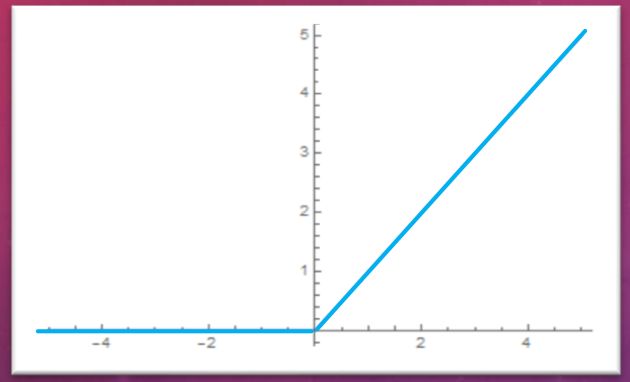
入力値×重み
を全て足す。
※重みwが学習するパラメータ

$$\begin{cases} a=x_1*w_1+x_2*w_2+x_3*w_3+b \\ y=f(a) \end{cases}$$

活性化関数にはいろいろある

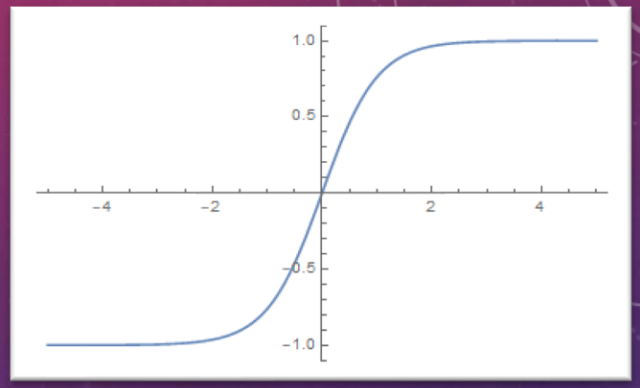
ReLU関数

$$F(x)=\begin{cases} x(x>0) \\ 0(x\leq 0) \end{cases}$$



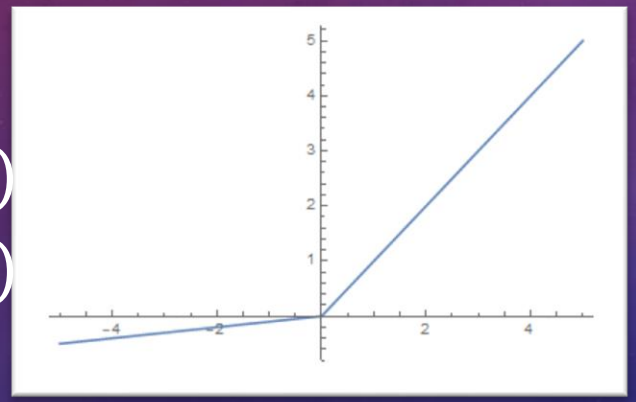
Tanh（ハイパボリックタンジェント）関数

$$F(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$$



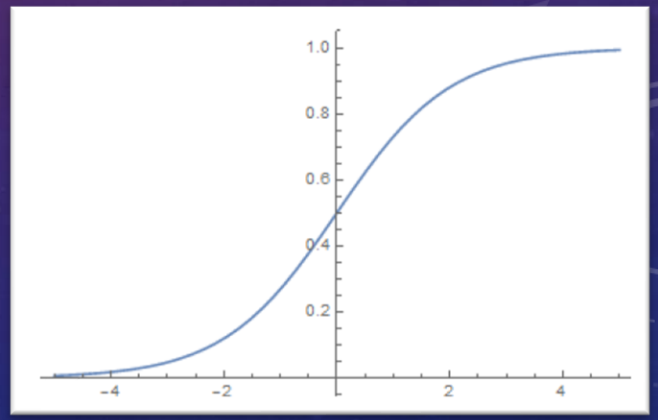
Leaky ReLU関数

$$F(x)=\begin{cases} x & (x>0) \\ 0.1x & (x\leq 0) \end{cases}$$



シグモイド関数

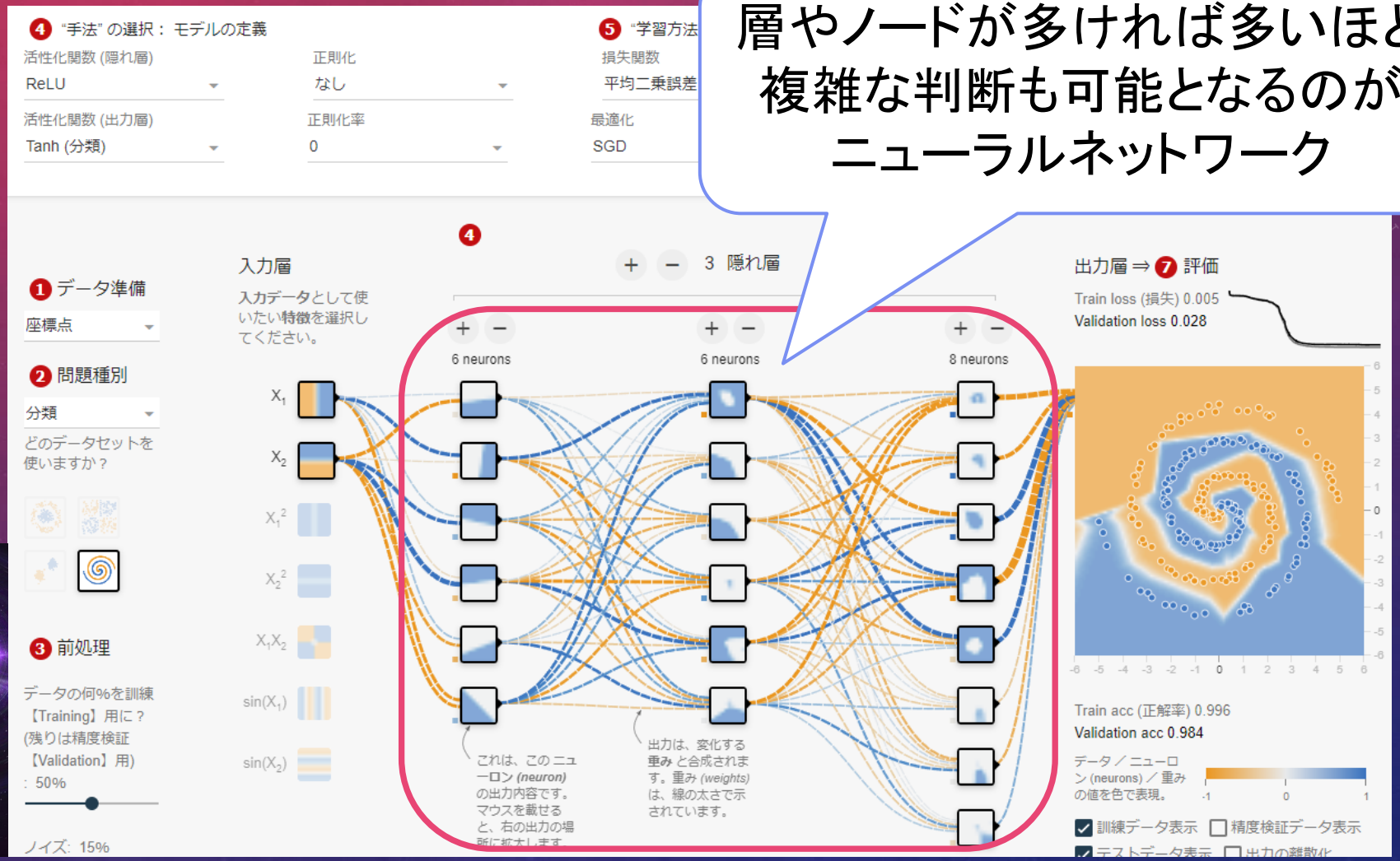
$$F(x)=\frac{1}{1+e^{-x}}$$



その他　ソフトマックス関数

3層でそれぞれ6、6、8ノードだと：2分類問題その３

層やノードが多ければ多いほど
複雑な判断も可能となるのが
ニューラルネットワーク

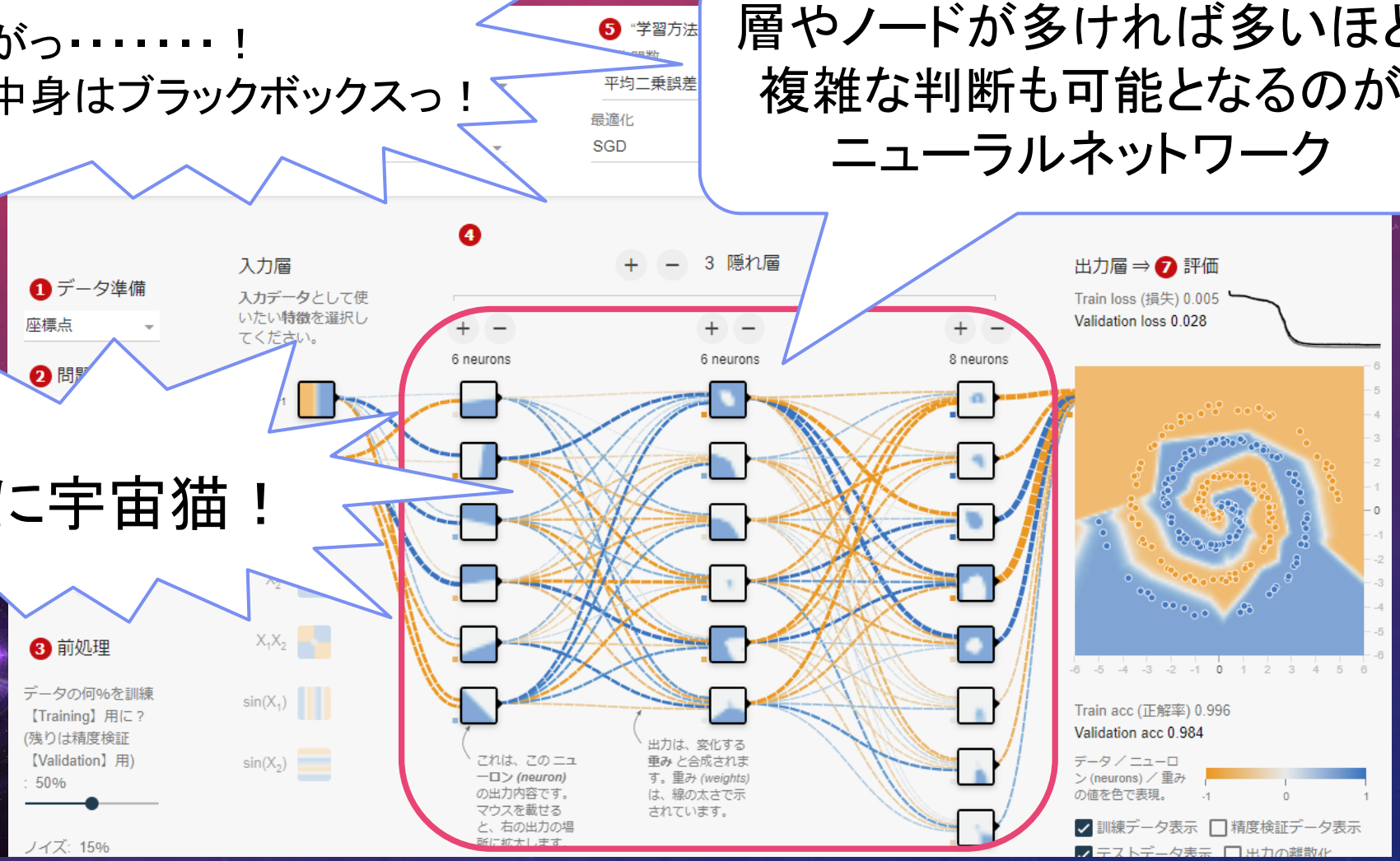


3層でそれぞれ6、8ノードだと：2分類問題その3

がっ……！
中身はブラックボックスっ！

層やノードが多ければ多いほど
複雑な判断も可能となるのが
ニューラルネットワーク

まさに宇宙猫！



ML Shukai

ニューラルネットワーク(NN)のタスク別設計

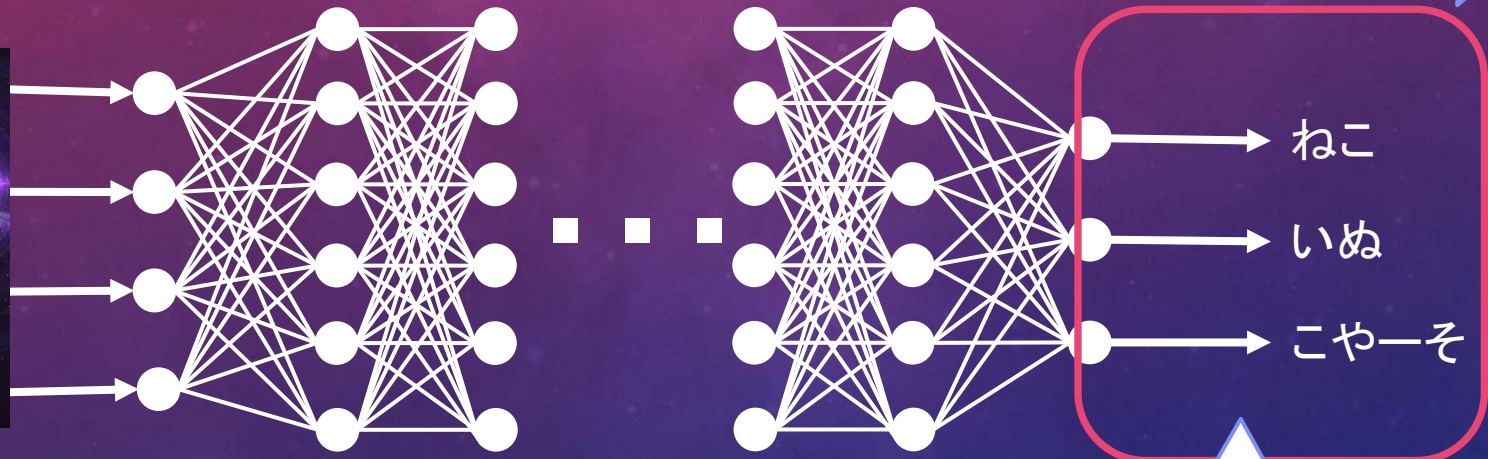
出力層の数と活性化関数と誤差関数

機械学習(ニューラルネットワーク含む)の大前提

モデルにデータを入れて
学習する

正解値と予測値から
線1本1本の重みを更新
(学習)

何度も学習した結果
誤差が小さくなったら
完成



誤差

入力データ
とセットの
答え

予測値

正解値

様々なタスクがAIによって出来ている

同じルールなのに
何でこんなに色々なこと
がAIに出来ているの？

AI何もわからん！

何処を設計するか？具体的には

実際には出力層の数と活性化関数、
誤差関数(誤差の求め方)を
うまく設計することで
様々なタスクに対応可能

ねこ
いぬ
こやーそ

予測値

誤差

入力データ
とセットの
答え

正解値

代表的なタスク

- 回帰（数値予測）
- 2分類問題（はい、いいえで答えられるような2つに分類できる問題）
- 多分類問題（3つ以上の種類に分類する問題）

回帰（数値予測）のタスク

- 数値予測の場合は**出力層は1個**
- **出力層に活性化関数は使用しない**（そのまま予測値として使う）≒恒等関数ともいう
- 誤差の求め方は**2乗和誤差**

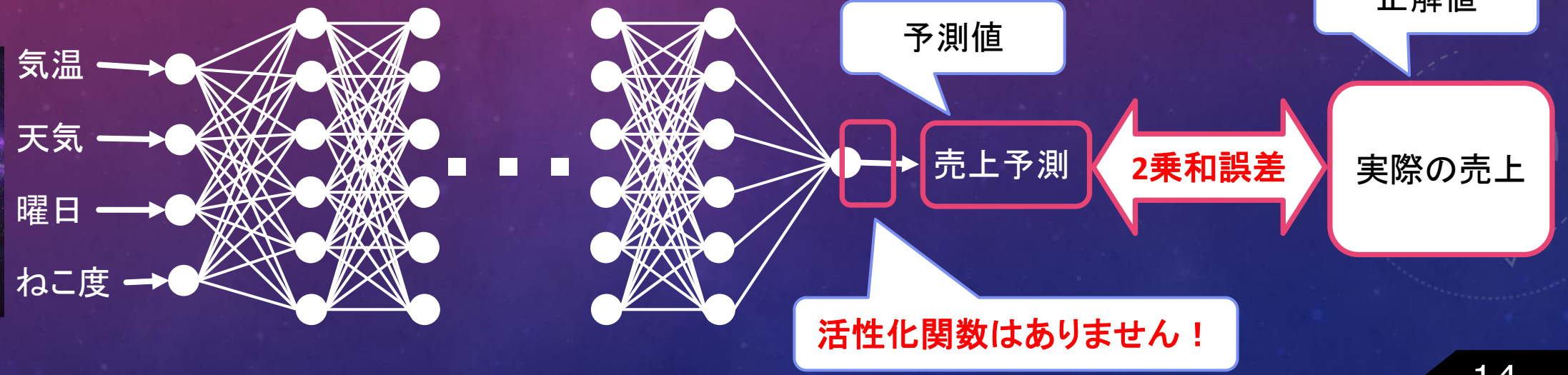
予測値：y'　正解値：y　$(y' - y)^2$

回帰（数値予測）のタスク

- 数値予測の場合は**出力層は1個**
- **出力層に活性化関数は使用しない**（そのまま予測値として使う）≒恒等関数ともいう
- 誤差の求め方は**2乗和誤差**

予測値：y'　正解値：y　$(y' - y)^2$

例：宇宙猫屋の売上予測



2分類問題のタスク

- 2分類予測の場合は**出力層は1個**
- 出力層の活性化関数は**シグモイド関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

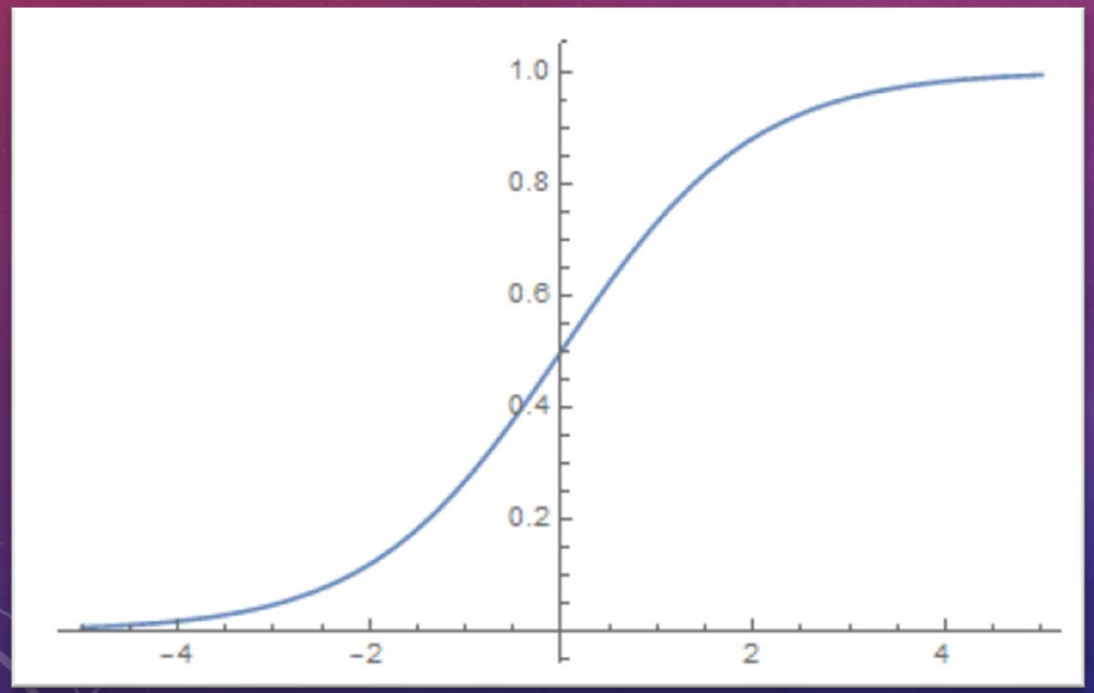
2分類問題のタスク

- 2分類予測の場合は**出力層は1個**
- 出力層の活性化関数は**シグモイド関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

2分類問題は「はい」か「いいえ」で
答えられる問題なので出力は1つでOK
また、計算できるよう正解は0と1で表現する。

2分類問題のタスク

- 2分類予測の場合は**出力層は1個**
- 出力層の活性化関数は**シグモイド関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

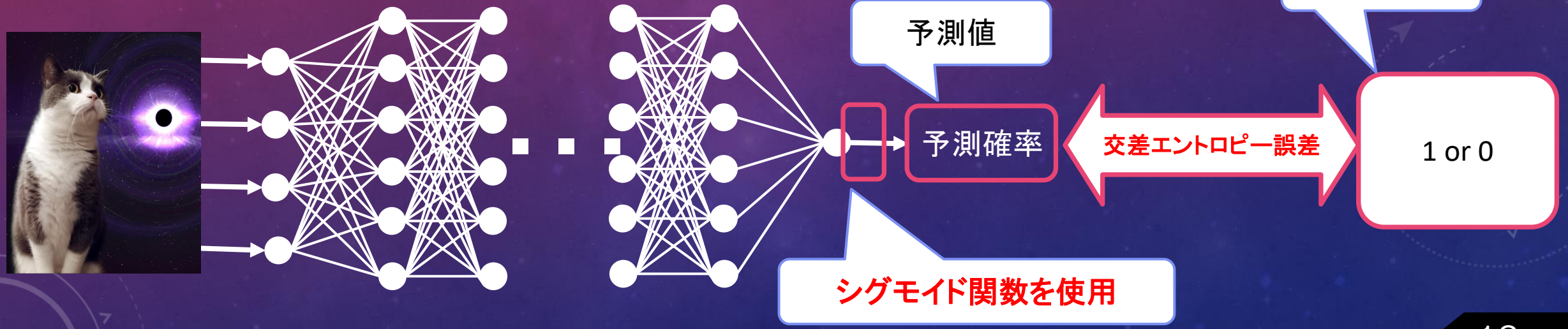


$F(x)=\frac{1}{1+e^{-x}}$ を通したものを出力
出力値は0～1の数値になる
（つまり、確率として扱える）

2分類問題のタスク

- 2分類予測の場合は**出力層は1個**
- 出力層の活性化関数は**シグモイド関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

例：画像が猫かどうか判定する（猫である＝1、猫でない＝0とする）

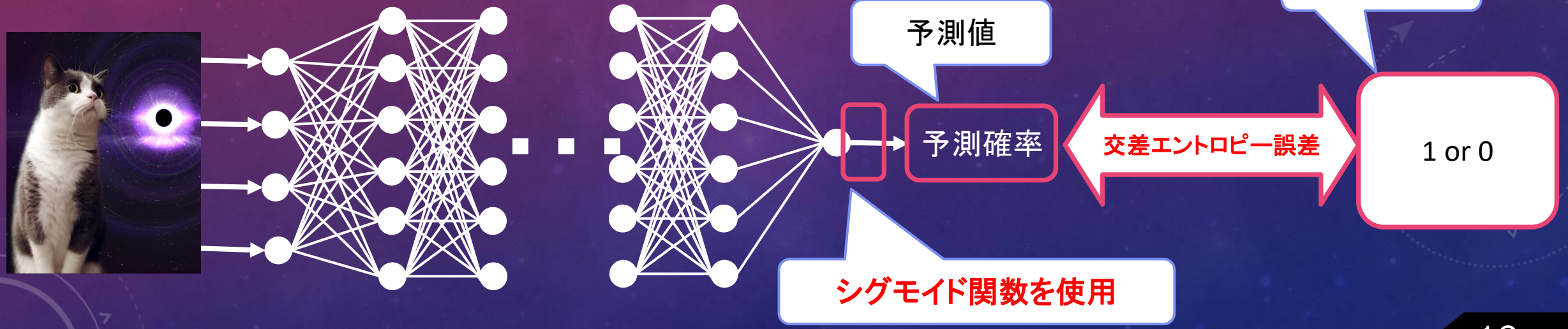


2分類問題のタスク

- 2分類予測の場合は**出力層は1個**
- 出力層の活性化関数は**シグモイド関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

最終出力は0～1の値なので
学習すれば
シグモイド関数を通した予測値は
予測の確率として考える事が出来る！

例：画像が猫かどうか判定する（猫である＝1、猫でない＝0とする）



2分類問題のタスク

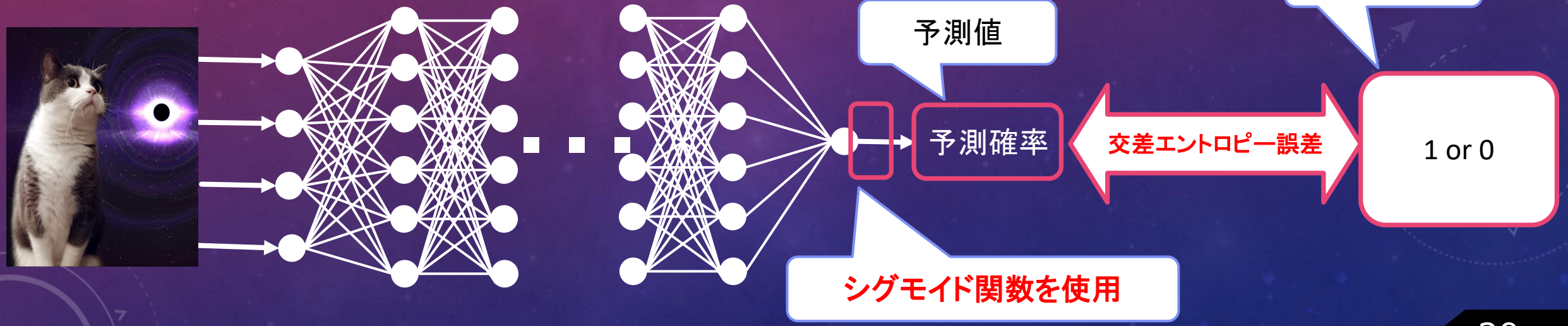
- 2分類予測の場合は**出力層は1個**
- 出力層の活性化関数は**シグモイド関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

交差エントロピー誤差
予測確率がどれだけ出にくい値なのかを表す計算式。

予測値：y'　正解値：y

$$-y \log y' - (1-y) \log(1-y')$$

例：画像が猫かどうか判定する（猫である＝1、猫でない＝0とする）



2分類問題のタスク

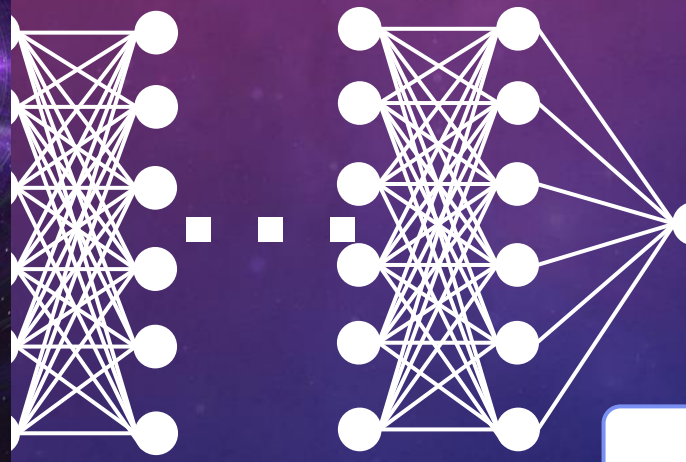
- 2分類予測の場合は**出力層は1個**
- 出力層の活性化関数は**シグモイド関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

交差エントロピー誤差
予測確率がどれだけ出にくい
値なのかを表す計算式。

$$-y \log y' - (1-y) \log(1-y')$$

？？？？？？？？？

する（猫である＝1、猫でない＝0とする）



予測値

予測確率

交差エントロピー誤差

正解値

1 or 0

シグモイド関数を使用

2分類問題

- 2
- 出力

を使用

交差エントロピー誤差
予測確率がどれだけ出にくい
値なのかを表す計算式。

$$-y \log y' - (1-y) \log(1-y')$$

猫でない＝0とする）

正解値

1 or 0

誤差

関数を使用

まさに宇宙猫！

だと思った？
残念！わたしです！

交差エントロピー誤差を簡単に

- $-y \log y' - (1-y) \log(1-y')$　　予測値：y'　　正解値：y
- 正解yが1のとき

$$-\log y'$$

- 正解yが0のとき

$$-\log(1-y')$$

- となります。
- どちらも「$-\log$(確率)」と考える事が出来るのが分かるでしょうか？
- これは情報量と言って確率が低ければ低いほど大きくなる値になります。

例：$y' = \frac{1}{2} = 0.5$のとき、$-\log_2 y' = 1$、　$y' = \frac{1}{4} = 0.25$のとき、$-\log_2 y' = 2$となります。

交差エントロピー誤差を簡単に

- $-y \log y' - (1-y) \log(1-y')$
- 正解yが1のとき

$$-\log y'$$

- 正解yが0のとき

$$-\log(1-y')$$

- となります。

正解が1のときy'が0に近づくほど
誤差$-\log_2 y'$は大きくなる。

正解が0のときはy'が1に近づくほど
誤差$-\log(1-y')$が大きくなる。

つまり、**正解から予測確率が離れれば離れるほど大きくなる値が交差エントロピー誤差。**

このため、誤差として計算しやすい計算式となっている。

- どちらも「$-\log$(確率)」と考える事が出来るのが分かるでしょうか？
- これは情報量と言って確率が低ければ低いほど大きくなる値になります。

例：$y' = \frac{1}{2}$のとき、$-\log_2 y' = 1$、　$y' = \frac{1}{4}$のとき、$-\log_2 y' = 2$となります。

交差エントロピー誤差を簡単に

- $-y \log y' - (1-y) \log(1-y')$
- 正解yが1[illegible]き

これ以上は難しいので

正解が1[illegible]き、
$-\log_2 y'$は、

正解が0のとき、
$-\log(1-y')$が[illegible]、

つまり、**正解から予測確率が離れれば離れるほど大きくなる値が情報量。**

のメ[illegible]誤差として計算しやすい
[illegible]っている。

計算式っ……！
覚えた方が早いっ！！

省　略

[illegible]確率[illegible]

[illegible]って確率[illegible]

[illegible]ょうか？

[illegible]になります。

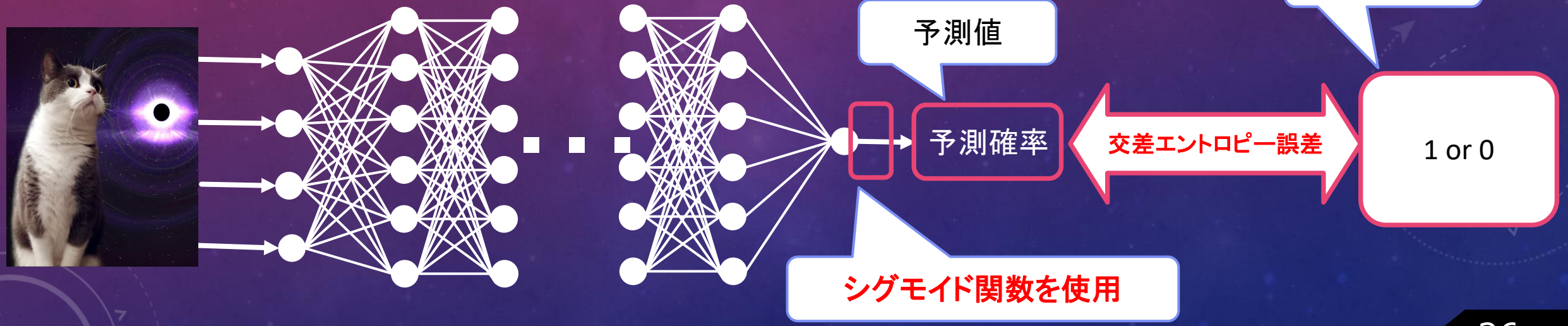
$\log_2 y' = 1$、[illegible]　$y' = \frac{1}{4}$のとき、$-\log_2 y' = 2$となります。

2分類問題のタスク

- 2分類予測の場合は**出力層は1個**
- 出力層の活性化関数は**シグモイド関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

予測値：y'　正解値：y　$-y \log y' - (1-y) \log(1-y')$

例：画像が猫かどうか判定する（猫である＝1、猫でない＝0とする）



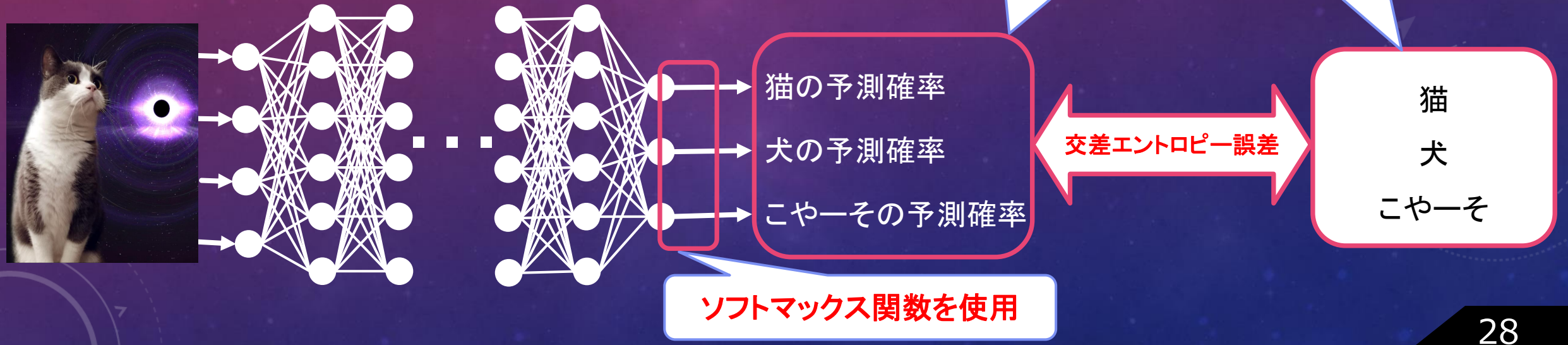
多分類問題のタスク

- 多分類予測の場合は**出力層はn個（nは予測したい分類数）**
- 出力層の活性化関数は**ソフトマックス関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

多分類問題のタスク

- 多分類予測の場合は**出力層はn個（nは予測したい分類数）**
- 出力層の活性化関数は**ソフトマックス関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

例：画像が猫かどうか判定する（「猫」「犬」「こやーそ」のどれかを予測）

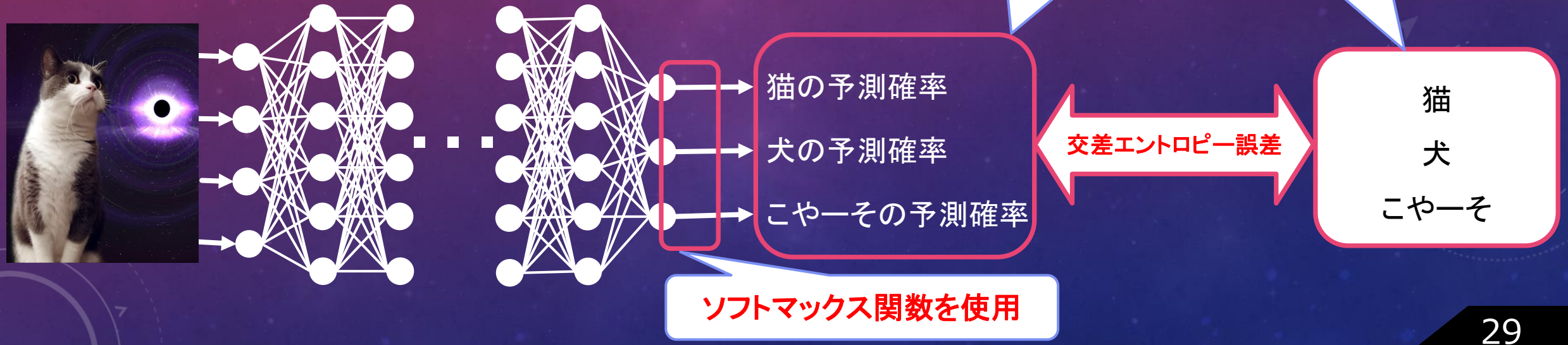


多分類問題のタスク

- 多分類予測の場合は**出力層はn個（nは予測したい分類数）**
- 出力層の活性化関数は**ソフトマックス関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

正解はどのように表すのか？

例：画像が猫かどうか判定する（「猫」「犬」「こやーそ」のどれかを予測）

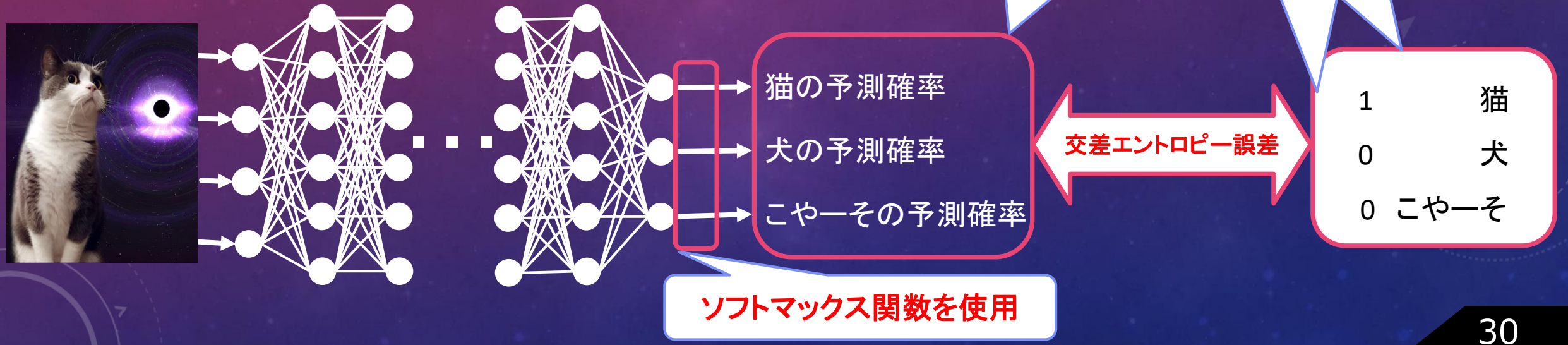


多分類問題のタスク

- 多分類予測の場合は**出力層はn個（nは予測したい分類数）**
- 出力層の活性化関数は**ソフトマックス関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

猫が正解の場合の正解値

例：画像が猫かどうか判定する（「猫」「犬」「こやーそ」のどれかを予測）

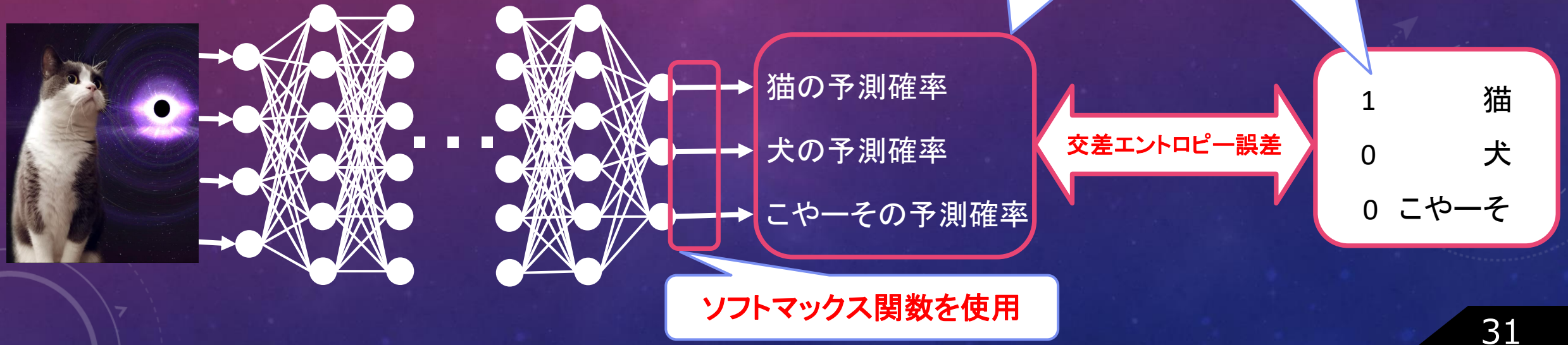


多分類問題のタスク

- 多分類予測の場合は**出力層はn個（nは予測したい分類数）**
- 出力層の活性化関数は**ソフトマックス関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

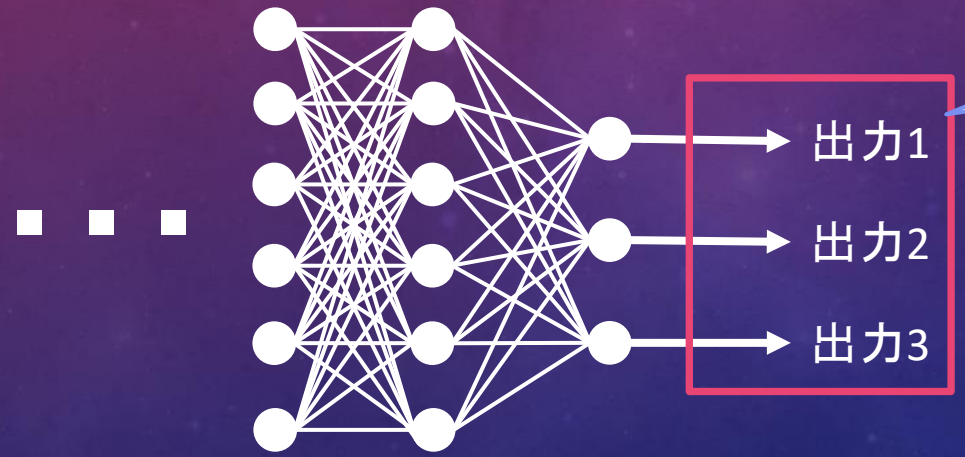
このように正解の選択肢を
0と1の分類個の数値で表す表現を
ワンホットベクトルと言います。

例：画像が猫かどうか判定する（「猫」「犬」「こやーそ」のどれかを予測）



ソフトマックス関数

- 複数分類のとき出力層で使われる関数
- 出力の合計値が1になる
- $[y_0, y_1, \cdots, y_{k-1}] = \frac{[e^{x0}, e^{x1}, \cdots, e^{xk-1}]}{e^{x0} + e^{x1} + \cdots + e^{xk-1}}$



出力の合計が1、つまり合計100%の確率として扱えるよ！

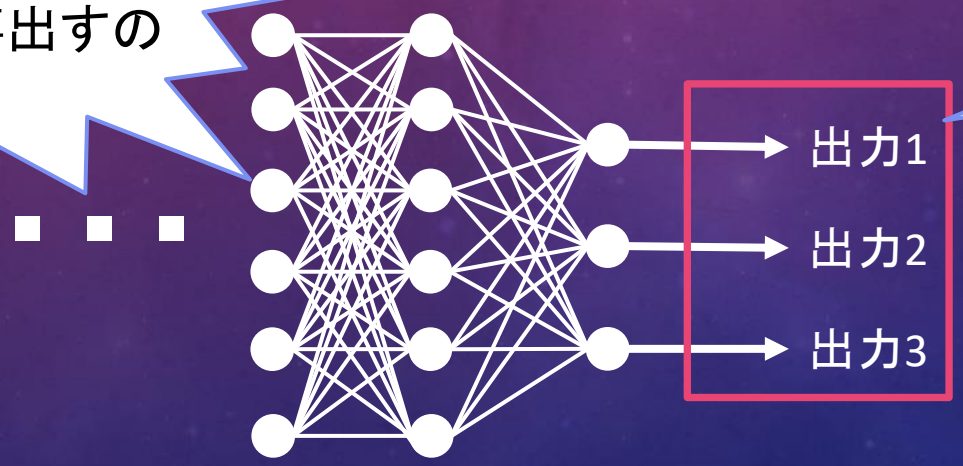
出力層が2個以上
つまり、複数分類問題
の出力層で使われるよ！

ソフトマックス関数

- 複数分類のとき出力層で使われる関数
- 出力の合計値が1になる
- $[y_0, y_1, \cdots, y_{k-1}] = \frac{[e^{x0}, e^{x1}, \cdots}{e^{x0} + e^{x1} + \cdots + e^{xk-1}}$

あーっ！お客様！
困ります！

あーっ！いけません！
分からない事出すの
いけません！



出力の合計が1、つまり合計
100%の確率として扱えるよ！

出力層が2個以上
つまり、複数分類問題
の出力層で使われるよ！

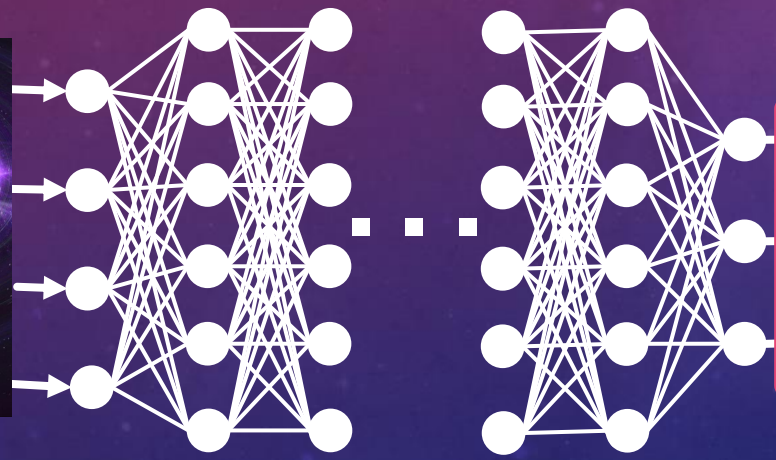
ソフトマックス関数(3分類の場合)

- 出力の合計値が1になる　それぞれの予測値をy'_0，y'_1，y'_2としたとき

$$y'_0=\frac{e^{x0}}{e^{x0}+e^{x1}+e^{x2}} \quad y'_1=\frac{e^{x1}}{e^{x0}+e^{x1}+e^{x2}} \quad y'_2=\frac{e^{x2}}{e^{x0}+e^{x1}e^{x2}}$$

$y'_0+y'_1+y'_2=1$ となります。

例：画像が猫かどうか判定する(「猫」「犬」「こやーそ」のどれかを予測)



予測値

y'_0猫の予測確率

y'_1犬の予測確率

y'_2こやーその予測確率

交差エントロピー誤差

猫が答えの場合の正解値

1	猫
0	犬
0	こやーそ

ソフトマックス関数(3分類の場合)

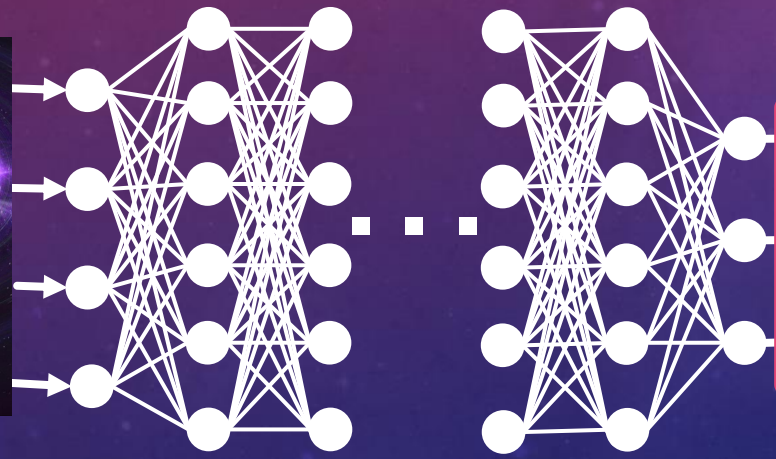
分母側はeの入力値の乗数を全て足したもの。
分子側はeの自身の入力値の乗数となる。

- 出力の合計値が1になる　それぞれの予測値を y'_0, y'_1, y'_2としたとき

$$y'_0 = \frac{e^{x0}}{e^{x0}+e^{x1}+e^{x2}} \quad y'_1 = \frac{e^{x1}}{e^{x0}+e^{x1}+e^{x2}} \quad y'_2 = \frac{e^{x2}}{e^{x0}+e^{x1}e^{x2}}$$

$y'_0 + y'_1 + y'_2 = 1$ となります。

例：画像が猫かどうか判定する(「猫」「犬」「こやーそ」のどれかを予測)



予測値

y'_0猫の予測確率
y'_1犬の予測確率
y'_2こやーその予測確率

交差エントロピー誤差

猫が答えの場合の正解値

1	猫
0	犬
0	こやーそ

ソフトマックス関数(3分類の場合)

学習が進めば一番高い数値が
可能性が高いと
判断できるようになる

- 出力の合計値が1になる　それぞれの予測値をy'_0 , y'_1 , y'_2としたとき

$$y'_0=\frac{e^{x0}}{e^{x0}+e^{x1}+e^{x2}} \quad y'_1=\frac{e^{x1}}{e^{x0}+e^{x1}+e^{x2}} \quad y'_2=\frac{e^{x2}}{e^{x0}+e^{x1}e^{x2}}$$

$y'_0+y'_1+y'_2=1$ となります。

例：画像が猫かどうか判定する(「猫」「犬」「こやーそ」のどれかを予測)

予測値

y'_0猫の予測確率

y'_1犬の予測確率

y'_2こやーその予測確率

交差エントロピー誤差

猫が答えの場合
の正解値

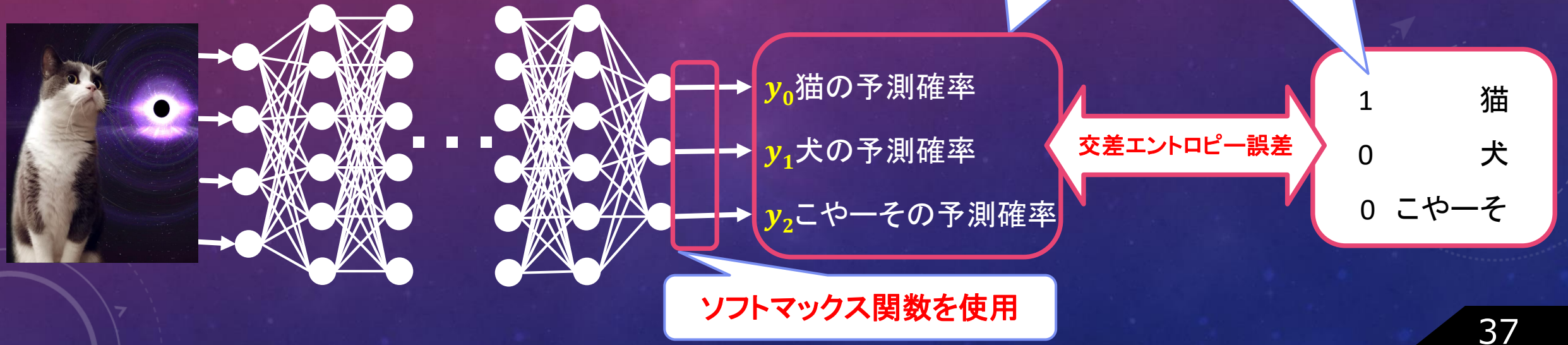
1	猫
0	犬
0	こやーそ

多分類問題のタスク

- 多分類予測の場合は**出力層はn個（nは予測したい分類数）**
- 出力層の活性化関数は**ソフトマックス関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

多分類の交差エントロピー誤差はどのような計算式？

例：画像が猫かどうか判定する（「猫」「犬」「こやーそ」のどれかを予測）



多分類の交差エントロピー誤差

- K分類（K>2)とした時 正解値をy、予測値をy'としたとき交差エントロピー誤差は以下のように表せられる。

$$-\sum_{k=0}^{K-1} y_k \, log \, y'_k$$

多分類の交差エントロピー誤差

- K分類（K>2)とした時 正解値をy、予測値をy'としたとき交差エントロピー誤差は以下のように表せられる。

$$-\sum_{k=0}^{K-1} y_k \, log \, y'_k$$

あー
そーゆーことね
完全に理解した

とでも言うと思ったか
なんもわからん！！

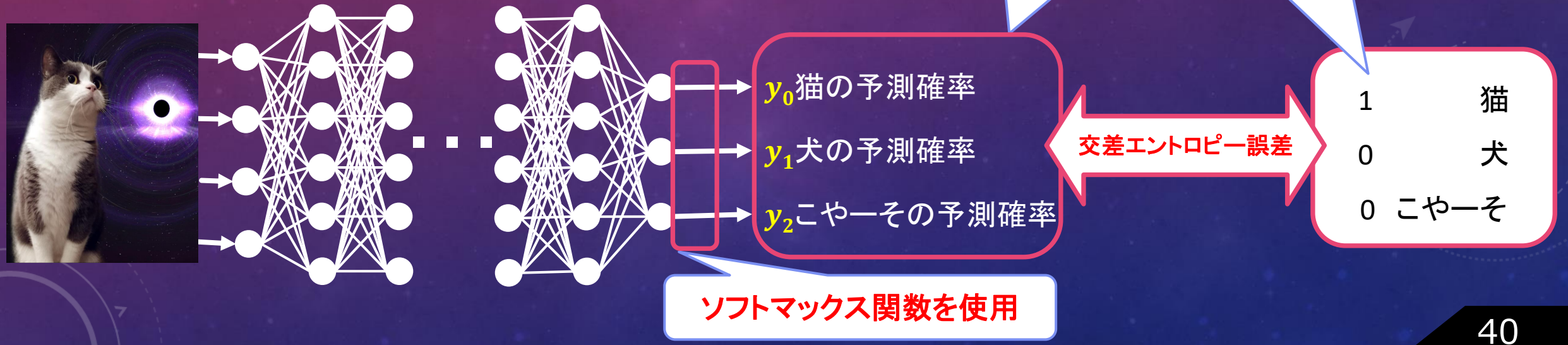
3分類の交差エントロピー誤差

- 予測値：y'　正解値：y　　猫が正解の場合　$y_0 = 1,\ y_1 = 0,\ y_2 = 0$ となるので代入すると

$$-\sum_{k=0}^{K-1} y_k \log y'_k = -(y_0 \log y'_0 + y_1 \log y'_1 + y_2 \log y'_2)$$

$$= -\log y'_0$$

例：画像が猫かどうか判定する（「猫」「犬」「こやーそ」のどれかを予測）



3分類の交差エントロピー誤差

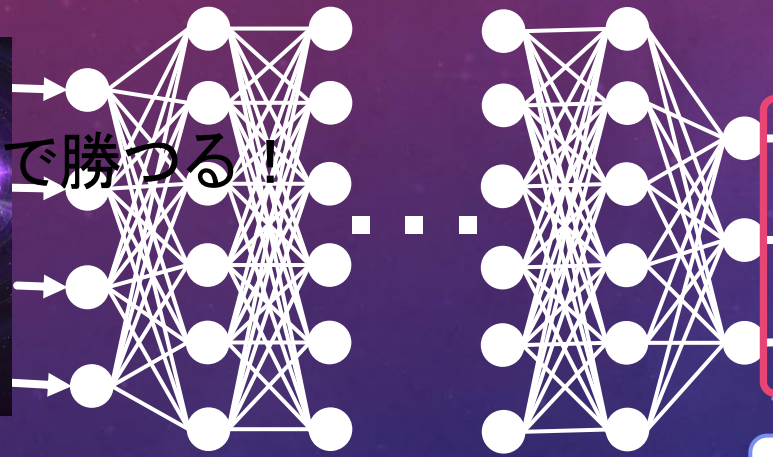
- 予測値：y'　正解値：y　　猫が正解の場合　$y_0 = 1, y_1 = 0, y_2 = 0$ となるので代入すると

$$-\sum_{k=0}^{K-1} y_k \log y'_k = -(y_0 \log y'_0 + y_1 \log y'_1 + \cdots + y_{K-1} \log y'_{K-1})$$

$$= -\log y'_0$$

2分類の時と同じ式になった！

例：画像が猫かどうか判定する（「猫」「犬」「こやーそ」のどれかを予測）



y_0猫の予測確率

y_1犬の予測確率

y_2こやーその予測確率

ソフトマックス関数を使用

予測値

交差エントロピー誤差

答えの場合の正解値

1	猫
0	犬
0	こやーそ

多分類問題のタスク

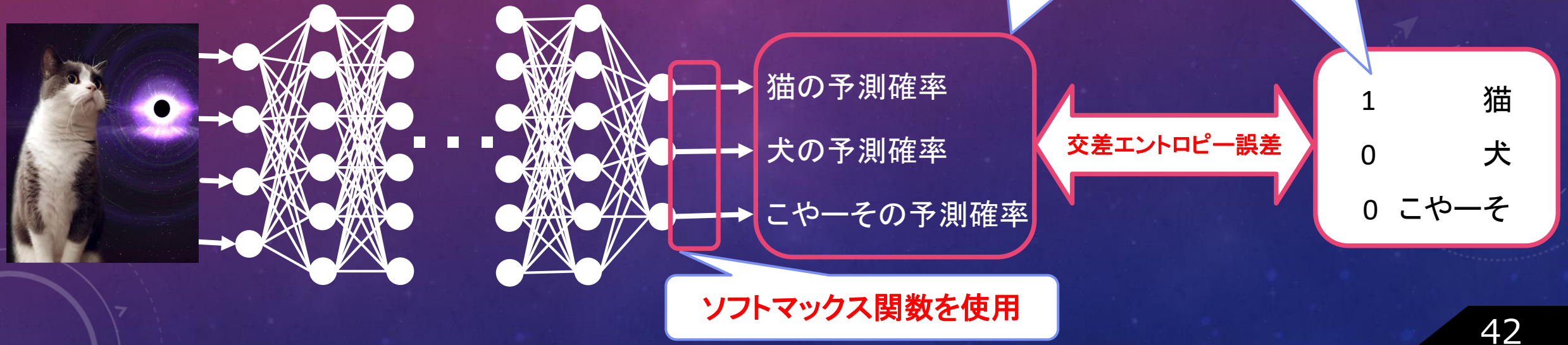
- 多分類予測の場合は**出力層はn個（nは予測したい分類数）**
- 活性化関数は**ソフトマックス関数**を使用
- 誤差の求め方は**交差エントロピー誤差**

予測値：y'　　正解値：y

$$-\sum_{k=0}^{K-1} y_k \, log \, y'_k$$

正解yの1がある所の
$-\log y'_k$
が誤差になる

例：画像が猫かどうか判定する（「猫」「犬」「こやーそ」のどれかを予測）



まとめ

- 基本的なAIは出力層の設計を変えることで様々なタスクに応用できる。
- 具体的には大きく分けて回帰（数値予測）、2分類問題、多分類問題の3種類がある。
- だいたいこれを使ったり、組み合わせたりすれば（データさえあれば）基本的な判断するAIが作れる。
  （ただし、画像生成・文章生成・強化学習などはさらに別の仕組みを追加する必要がある。）

- **回帰**は出力層は**1個**、**活性化関数無し**
- **2分類**は出力層は**1個**、活性化関数は**シグモイド関数**
- **多分類**は出力層は**n個**、活性化関数は**ソフトマックス関数**

- **y=ax+bは流石に今回は関係なかった…**

おわり

ありがとうございました

第3回　y=ax+bから始める初心者向けML講座

次回は最適化法（確率的勾配降下法の応用）、または画像を入力するには？のどちらか（多分）