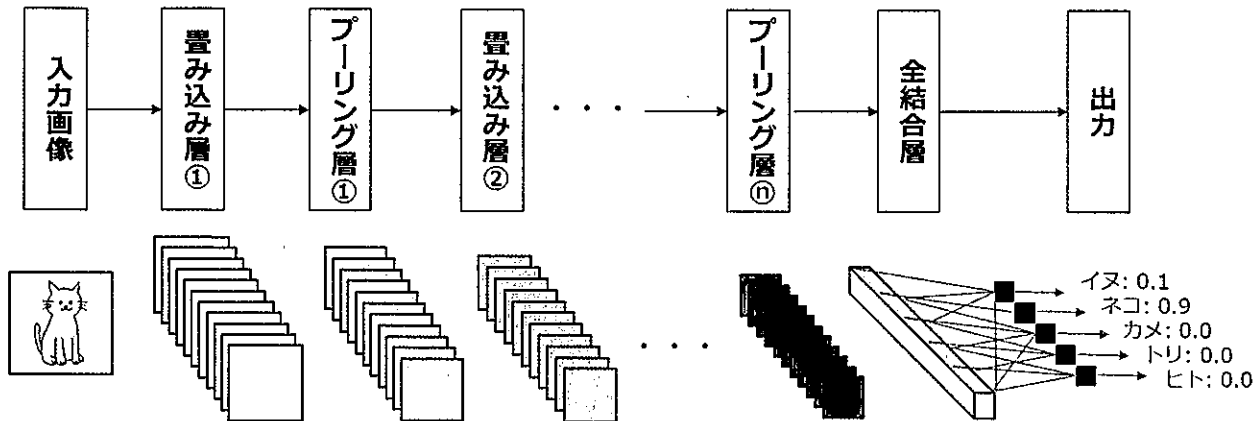


電子情報 (データサイエンス)

問題 1. 以下の問いについて、それぞれ指定された答案用紙に解答しなさい。



上図の畳み込みニューラルネットワーク (CNN) でネコの画像が入力され、畳み込み層①、プーリング層①、畳み込み層②、 \dots 、プーリング層②の順に信号が伝搬していき、全結合層まで到達すると、以下の式に基づいて、出力 y_j ($j = 1, \dots, J$) が求められるとする。

$$y_j = f(u_j) = \frac{\exp(u_j)}{\sum_{k=1}^J \exp(u_k)} \quad \text{但し, } u_j = \sum_{i=1}^I w_{ji}x_i + b_j$$

ここで、 I, J は、それぞれ全結合層の入力数と出力数である。 x_i, w_{ji}, b_j はそれぞれ、全結合層 i 番目の入力値、 i 番目の入力から j 番目の出力への結合荷重、 j 番目のバイアス値を表す。また、関数 $f(\cdot)$ はソフトマックス関数と呼ばれ、出力 y_j を入力画像がクラス j である確率として解釈したいときに使われる。

- (1) 関数 $f(u_j)$ の微分を出力 y_j で表しなさい。
- (2) 上図の CNN で入力画像ごとにクラス情報 (上図では、イヌ、ネコ、カメ、トリ、ヒトがクラスに対応) が与えられ、それを教師信号 $\vec{d} = (d_1, d_2, \dots, d_J)$ で表す。このとき、出力と教師信号の二乗誤差を定義し、最急降下法に基づいて、結合荷重 w_{ji} の更新式を導出しなさい。
- (3) 上図の CNN において、畳み込み層①からプーリング層②に至るまでの処理は、どのような機能を実現していると言えるかを 300 字以内で説明しなさい。説明において、畳み込み層やプーリング層の特性、そして、これらを繰り返して配置する理由が述べられているものほど、高い評価を与える。

問題 2. 以下の問いについて、それぞれ指定された答案用紙に解答しなさい。

- (1) SNS へのスパム投稿対策のため、あるスパムフィルターの性能評価を行いたい。このスパムフィルターは、スパム投稿を 98%の確率で正しく検知できるが、正常投稿を 2%の確率で誤って判定する。ある投稿がこのスパムフィルターで「スパム」と判定されたとき、この投稿が実際にスパムである確率をベイズの定理を用いて計算しなさい。なお、この SNS におけるスパム投稿の割合は、全投稿の 5%とする。
- (2) ある SNS では、全投稿の 5%がスパムとする。また、正常投稿とスパム投稿には、「ビットコイン」、「クジラ」、「買い増し」というトークンが下表の確率で出現する。

トークン	トークン出現率	
	正常投稿	スパム投稿
ビットコイン	0.2	0.8
クジラ	0.05	0.7
買い増し	0.1	0.5

このとき、「ビットコイン」、「クジラ」、「買い増し」を含む投稿が正常かスパムかを、単純ベイズ分類器により判定しなさい。ここで、投稿に含まれるトークンの出現確率は互いに独立であるとする。

- (3) スパムフィルターを構成する方法を一つ取り上げ、300 字以内で説明しなさい。説明において、使用する分類器モデルと文書特徴量の一例を取り上げ、それらの特性をより適切に説明しているものほど高い評価を与える。