

大連理工大学

姓名 _____
学籍番号 _____
学部 _____
クラス _____
科目番号 _____
担当教員: 福森隆寛

科目名: 機械学習 問題種類: A 試験形式: 閉 卷
所属学部: 国際情報ソフトウェア学部 試験時間: _____
問題用紙合計ページ数 4

	I	II	III	IV							合計
配点	60	15	12	13							100
得点											

装

訂

線

[問題 I] に入る言葉を、選択肢から 1 つ選択せよ。

(※ 同じ番号の には、同じ言葉が入る。)

- 正解の付いたデータを用いた学習を教師あり学習と言う。教師あり学習には、入力を予め定められたクラスのいずれかに分類する (1) や、入力から予想される妥当な出力値を求める (2) がある。 (1) に関する手法の 1 つとして、与えられる個々のデータからあるクラスについて共通点を求める (3) がある。 (3) の代表例として、データを分類する質問をノードとし、分類結果をリーフとする木構造を用いる (4) がある。 (4) の問題点は、木が成長しすぎて学習データに適応しすぎる (5) という現象が発生しやすいことである。

(1) [A] 識別、[B] モデル推定、[C] 強化学習、[D] ベイズの定理

(2) [A] データマイニング、[B] サポートベクトルマシン、
[C] 交差確認法、[D] 回帰

(3) [A] ベイジアンネットワーク、[B] 概念学習、
[C] ナイーブベイズ識別、[D] 最尤推定法

(4) [A] 決定木、[B] 最小二乗法、[C] ロジスティック識別、[D] カーネル関数

(5) [A] 勾配消失問題、[B] ゼロ頻度問題、[C] 過学習、[D] マルコフ性

- 正解が付いていないデータを用いた学習を教師なし学習と言う。教師なし学習には、入力データから何らかの共通点をもつデータをまとめる (6) や、大量のデータから何度も出現するパターンなどを発見する (7) がある。 (6) の代表的な手法として、分類対象の集合を内的結合と外的分離が達成されるような部分集合に分割する (8) という操作がある。 (8) の結果を用いて、新たに観測されたデータが属するクラスを決めることを考えたとき、各クラスタの確率分布 (確率密度関数) を仮定して、そのパラメータを (9) という手法を使って推定する。 (7) の代表的な手法として、a priori な原理の対偶を用いて、膨大な項目集合や連想規則を絞り込む Apriori アルゴリズムが提案されており、これを高速化した手法が (10) である。

- (6) [A] ID3 アルゴリズム、[B] モデル推定、[C] 誤り訂正学習、[D] 線形回帰
 (7) [A] データマイニング、[B] 回帰、[C] ソフトマージン、[D] AdaBoost
 (8) [A] クラスタリング、[B] 正則化、[C] 異常検出、[D] 連想規則抽出
 (9) [A] ビタビアルゴリズム、[B] YATSI アルゴリズム、
 [C] Value Iteration アルゴリズム、[D] EM アルゴリズム
 (10) [A] FP-Growth アルゴリズム、[B] FIND-S アルゴリズム、
 [C] ID3 アルゴリズム、[D] 候補削除アルゴリズム

- 教師あり学習や教師なし学習のいずれにも該当しない手法として、正解付きデータから得られた識別器の性能を正解無しデータを使って向上させる (11) や、最も多くの報酬を得られるように各状態における最適な出力を獲得させる (12) などがある。(11) の手法の1つとして、異なった特徴を用いて識別器を2つ作成し、相手の識別結果を利用しながら、それぞれの識別器を学習させる (13) などがある。(12) はQ値(ある行為によって得られる報酬の推定値)が大きくなるような行為を選択するが、状態遷移確率や報酬の確率分布が与えられる (14) のような場合、動的計画法を用いてQ値を計算する。一方、このような確率情報が与えられないモデルフリーである場合、(15) のような手法で最適な意思決定規則を学習する。

- (11) [A] 概念学習、[B] 過学習、[C] 半教師あり学習、[D] 事前学習
 (12) [A] 回帰木、[B] AdaGrad、[C] 強化学習、[D] 最急勾配法
 (13) [A] 自己学習、[B] ラベル伝搬法、[C] YATSI アルゴリズム、[D] 共訓練
 (14) [A] 生成モデル、[B] モデルベース、[C] 識別モデル、[D] モデル推定
 (15) [A] TD 学習、[B] 誤り訂正学習、[C] アンサンブル学習、[D] 表現学習

- 機械学習は、「データ収集 → (16) → 評価基準の設定 → 学習 → 結果の可視化」の順番で処理される。(16) は収集したデータに何かしら手を加えることであり、主成分分析によって特徴ベクトルの次元数を削減したり、(17) によって特徴の値を揃えたりする。結果の可視化は、識別結果からいくつかの評価指標を計算し、表やグラフとして表示することである。特に識別問題では4種類の要素(True Positive、False Negative、False Positive、True Negative)で構成された混同行列から、正解率、精度、(18)、(19)などを計算することが多い。(18) は正例が正しく判定された割合を表し、(19) は精度と(18)の調和平均を表す。2クラスの分類問題では、FPR (False Positive Rate) とTPR (True Positive Rate) の関係を表す (20) に基づいて、観測された特徴を分類するための適切な閾値を設定する。

- (16) [A] m 推定、[B] カーネルトリック、[C] 表現学習、[D] 前処理
 (17) [A] 標準化、[B] 枝刈り、[C] 欲張り法、[D] 識別関数法
 (18) [A] 活性化関数、[B] エントロピー、[C] 局所異常因子、[D] 再現率
 (19) [A] 損失関数、[B] バイアス、[C] F 値、[D] 分散
 (20) [A] ROC 曲線、[B] CART、[C] Boosting、[D] CRF

錯
対
錯
錯
対

[問題Ⅱ] 次の文章で読んで、正しい説明には [A] を、誤っている説明には [B] を選択せよ。

margin只是代表距离

- (21) マージンとは、識別面と最も近い学習データとの距離が最大となる識別超平面を求める手法である。SVM
- (22) ニューラルネットワークのノード同士を結合する重みは、誤差逆伝播法などを用いて最適化されるが、重みの修正量が入力層の方向へ戻につれて小さくなる勾配消失問題が発生することがある。
- (23) 畳み込みニューラルネットワークは、畳み込み層、プーリング層、LSTMセルを配置して、最後に全結合のニューラルネットワークを配置した構造を有する。
- (24) 誤りを減らすことに特化した識別器を次々に加える方法で異なる振る舞いをする識別器の集合を作成するランダムフォレストは、アンサンブル学習の1つである。
不是随机树，是boosting
- (25) 学習データが特徴空間上で線形分離可能なら、パーセプトロンの学習アルゴリズムで線形分離面を見つけることが出来る。

[問題Ⅲ] 1次元の特徴量 x を用いて、入力データを正例、負例のクラスのどちらかに識別する問題を考える。表1に示す4つの学習データが与えられたとき、パーセプトロンの学習アルゴリズムを適用して、識別関数の重みを学習する。このときの学習係数は、 $\eta = 2.0$ とする。識別関数は、 $g(x) = w_0 + w_1x$ で表現され、 $g(x) > 0$ であれば正例、 $g(x) < 0$ であれば負例とする。重みの初期値は、 $w_0 = 0.5$ 、 $w_1 = 1.0$ とする。

表1 学習データ

番号	特徴 x	クラス
1	1.2	正例
2	-0.2	正例
3	-0.4	負例
4	-0.8	負例

- (26) 番号1から順にパーセプトロンの学習アルゴリズムを適用するとき、最初に重み係数が更新されるデータを以下の選択肢から選択せよ。

[A] 番号1、[B] 番号2、[C] 番号3、[D] 番号4

- (27) パーセプトロンの学習アルゴリズムによって最終的に決定される重みを以下の選択肢から選択せよ。

[A] $(w_0, w_1) = (-1.5, 1.8)$ 、[B] $(w_0, w_1) = (0.5, 1.4)$ 、
[C] $(w_0, w_1) = (1.8, -1.5)$ 、[D] $(w_0, w_1) = (1.4, 0.5)$

- (28) この学習によって決まる正例クラスと負例クラスを分離する境界 ($g(x) = 0$) を以下の選択肢から選択せよ。

[A] $x = 1.8/1.5$ 、[B] $x = 1.5/1.8$ 、[C] $x = -1.4/0.5$ 、[D] $x = -0.5/1.4$

【問題Ⅳ】 図1のように、クラス1(C_1)とクラス2(C_2)のHMMが与えられたとする。

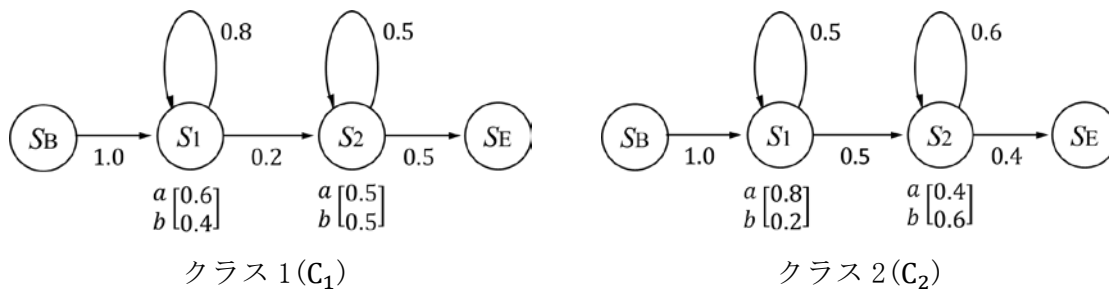


図1 HMM

(29) C_1 のHMMから特徴ベクトル系列 $\mathbf{x} = \text{"baa"}$ が出力される確率 $P(\mathbf{x}|C_1)$ を
 C 以下の選択肢から選択せよ。

[A] 0.005、[B] 0.0096、[C] 0.0146、[D] [A]～[C]以外

(30) C_2 のHMMから特徴ベクトル系列 $\mathbf{x} = \text{"baa"}$ が出力される確率 $P(\mathbf{x}|C_2)$ を
 B 以下の選択肢から選択せよ。

[A] 0.0064、[B] 0.01024、[C] 0.0384、[D] [A]～[C]以外

(31) C_1 と C_2 の事前確率が $P(C_1) = 0.3$ 、 $P(C_2) = 0.7$ のとき、特徴ベクトル系列 $\mathbf{x} = \text{"baa"}$ が分類されるクラスを以下の選択肢から選択せよ（言い換えると、 $P(C_1|\mathbf{x})$ と $P(C_2|\mathbf{x})$ を計算して、これら2つの事後確率の内、大きい方のクラスを選択せよ）。

[A] クラス1、[B] クラス2、[C] 両方のクラス、[D] どちらにも分類されない