

0.1. Карасева Т.С., Мамонтов Д.Ю. Сравнение эффективности методов интеллектуального анализа данных при решении задач банковской сферы.

В настоящее время все более актуальным становится применение методов интеллектуального анализа данных в банковской сфере. В ходе данной работы были протестированы методы классификации при решении пяти различных задач данной области. Анализ данных осуществлялся с помощью системы RapidMiner [1]. Были использованы алгоритмы: наивный байесовский классификатор (NB); индукция правил (IR); метод ближайших соседей (k-NN); дерево решений (DT); нейронная сеть (ANN); линейный дискриминантный анализ (LDA); случайный лес (RF). Проведено сравнение используемых методов на наличие статистически значимого различия. Далее указаны методы показавшие лучшую (Best) и худшую (Worst) эффективность в каждой из задач.

Задача №1 Best: DT, IR, NB; Worst: LDA.

Задача №2 Best: DT, IR; Worst: ANN.

Задача №3 Best: DT, IR, ANN, k-NN; Worst: LDA.

Задача №4 Best: k-NN; Worst: LDA.

Задача №5 Best: DT, NB, IR, ANN, LDA, RF; Worst: k-NN. Далее были реализованы два типа ансамблей.

Первый: содержит методы, показавшие наилучшую точность, второй: два метода, показавшие наибольшую точность, и метод с наименьшим результатом вектора производительности. Использование оператора ансамблирования позволило увеличить точность классификации. Последние мета-методы, которые были опробованы, это Баггинг (BG) (вид ансамбля для одного алгоритма) и Бустинг (BT). Далее приведены результаты для всех поставленных задач. В скобках указаны алгоритмы для которых применялись выбранные мета-методы. Задача №1. BG (DT): 83,72%; BT (DT): 83,45%. Задача №2. BG (IR): 92,57%; BT (IR): 89,06%. Задача №3. BG (DT): 93,17%; BT (DT): 94,07%. Задача №4. BG (k-NN): 93,80%; BT (k-NN): 100%. Задача №5. BG (LDA): 76,50%; BT (LDA): 77,00%. Можно сделать вывод, что для различных задач банковской сферы эффективными оказываются различные методы, а используемая система RapidMiner не обеспечивает достаточную точность, поэтому следует использовать другие системы, содержащие, такие подходы, как генетическое программирование [2], в том числе самонастраивающиеся методы [3].

[2] СЕМЕНКИН Е.С., ШАБАЛОВ А.А., ЕФИМОВ С.Н. Автоматизированное проектирование коллективов интеллектуальных информационных технологий методом генетического программирования // Вестник Сибирского государственного аэрокосмического университета им. академика М.Ф. Решетнева. — 2011. — №3 (36), С. 77–81.

[3] SEMENKIN E., SEMENKINA M. Self-configuring genetic programming algorithm with modified uniform crossover // 2012 IEEE Congress on Evolutionary Computation, CEC 2012, — 2012. — С. 6256587.

Список литературы

- [1] RapidMiner [Электронный ресурс]. URL: <https://rapidminer.com/> (дата обращения: 11.09.2015).