

ФОРМИРОВАНИЕ КОЛЛЕКТИВА РЕШАЮЩИХ ПРАВИЛ МНОГОКРИТЕРИАЛЬНЫМ ЭВОЛЮЦИОННЫМ АЛГОРИТМОМ В ЗАДАЧЕ АНАЛИЗА ЭМОЦИЙ ЧЕЛОВЕКА ПО АУДИОДАНЫМ

А.С. Полякова
Л.В. Липинский

polyakova_nasty@mail.ru
lipinskiyl@mail.ru

Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, г. Красноярск, Российская Федерация

Аннотация

Одной из важнейших задач современного этапа информатизации общества является развитие систем человеко-машинного интерфейса, в том числе систем автоматизированного распознавания эмоций человека. Эмоции человека выражены в виде комбинации двух показателей: Valence — направленность эмоции (отрицательные или положительные) и Arousal — выраженность эмоции (степень возбужденности). Эти показатели представлены вещественными числами. Для точности прогнозирования предложено использовать методы коллективного вывода. Точность коллективного решения оценена коэффициентом корреляции согласованности. Агенты (алгоритмы) в коллектив отобраны с помощью многокритериального эволюционного алгоритма. Применение многокритериального эволюционного алгоритма дало возможность автоматизировать процесс формирования коллектива, что позволило снизить затраты временных и материальных ресурсов. Коллектив построен на основе двух критериев: максимизации точности и минимизации числа агентов в коллективе. Для принятия решения коллективом применены такие способы, как простое голосование, взвешенное среднее, взвешенное среднее с учетом степени доверия к агенту, а также система на основе нечеткой логики. Для повышения эффективности решения задачи интеллектуального анализа данных предложена модификация системы на основе нечеткой логики. Проведены анализ и исследование эффективности многокритериального эволюционного алгоритма при решении задачи

Ключевые слова

Коллектив, эволюционные процедуры, многокритериальные алгоритмы оптимизации, алгоритмы интеллектуального анализа данных, системы на нечеткой логике, задача регрессии

прогнозирования эмоционального поведения человека. Экспериментально выявлено, что применение многокритериального эволюционного алгоритма для автоматизированного формирования состава коллектива приводит к более точному решению задачи

Поступила 15.01.2019
© Автор(ы), 2019

Исследование выполнено при поддержке Министерства науки и высшего образования Российской Федерации (Минобрнауки России) в рамках базового госбюджетного финансирования по проекту № 2.1680.2017/ПЧ

Введение. Для повышения точности решения задач классификации и регрессии в последние годы широко используются коллективы моделей. В литературе подобные системы упоминаются под разными названиями: смесь экспертов (mixture of experts), ансамбли классификаторов (classifier ensembles), смесь классификаторов (classifier fusion), системы множественных классификаторов (multiple classifier systems) [1]. Ансамбль классификаторов (коллектив) — это набор классификаторов, чьи индивидуальные решения каким-то образом объединяются (обычно путем голосования) для классификации новых примеров. Хансен Л.К. и Саламон П. в 1990 г. показали, что ансамбль классификаторов может быть более эффективным, чем один, если каждый классификатор ансамбля отличается от других с точки зрения ошибки классификации [2]. Это означает, что одной из основных проблем в объединении классификаторов является «создание разнообразных классификаторов».

Для построения коллективных моделей алгоритмов могут быть реализованы различные типы стратегий. Одна из стратегий состоит в том, что необходимо изменить параметры алгоритмов в коллективе. Это позволит внести разнообразие в процесс обучения, чтобы можно было получить различные модели обучения (шаблоны, паттерны). Эта идея была приведена в работе [3]. Другой популярный подход к построению коллективов заключается в манипулировании примерами обучения. Общая техника известна как бустинг [4] — когда для создания агентов в коллективе используются различные случайно выбранные подмножества. Однако вероятность выбора данных варьируется в зависимости от результатов, полученных из ранее созданных агентов.

Если объединить (агрегировать) результаты предсказания различных агентов, обученных на случайных подмножествах (элементы которых могут пересекаться или дублироваться), то такой метод называется бэггинг [5].

Поскольку успех коллективов алгоритмов зависит от набора соответствующих агентов и их разнообразия, возникает необходимость разработки процедуры автоматизированного отбора агентов в состав коллектива. Избыток агентов приводит к возникновению эффекта переобучения. Недостаток агентов может не привести к повышению точности решения задачи. В связи с этим возникает необходимость учета нескольких критериев.

В настоящей работе для учета критериев предлагается применять многокритериальные эволюционные алгоритмы (МЭА) NSGA-II [6] и SPEA-II [7]. Эти алгоритмы позволят автоматизировать формирование состава коллектива, тем самым экономя вычислительные ресурсы (позволят минимизировать число агентов) и решая поставленные задачи достаточно качественно (повышая способность к обобщению результата).

Алгоритм коллективного принятия решения на основе нечеткой логики. Задача распознавания эмоций является задачей классификации, для решения которой используют традиционные классификаторы, такие как смесь гауссовых распределений (GMM), скрытые марковские модели (HMM), искусственные нейронные сети (ANN), метод опорных векторов (SVM), обобщенно-регрессионные нейронные сети (GRNN), глубокие нейронные сети (DNN).

В последние годы все больше внимания уделяется автоматическому и непрерывному прогнозированию эмоций спонтанного поведения людей [8]. Для такой задачи были предложены и исследованы различные регрессионные модели: регрессия опорных векторов (SVR) [9], регрессия вектора релевантности (RVR) [10], нейронные сети прямой проводимости (FNNs) [11], рекуррентные нейронные сети (RNN) [12] и др. При использовании разных размеров одного признакового пространства для непрерывного прогнозирования возбуждения и валентности в различных базах данных выявлено, что одни алгоритмы дают лучше результат при использовании наименьшего признакового пространства, а другие алгоритмы — на всем признаковом пространстве.

Например, в работе [13] сделан вывод о превосходстве алгоритма SVR над BLSTM-RNN, а в работе [14] наоборот, о превосходстве второго метода. Разумным объяснением этому может быть то, что каждая модель прогнозирования имеет свои плюсы и минусы. Например, BLSTM-RNN очень чувствительны к переоснащению, но SVR не может явно моделировать контекстные зависимости.

Многие алгоритмы машинного обучения были применены к проблеме прогнозирования эмоционального состояния человека. Для объединения преимуществ различных регрессионных моделей используют раз-

личные коллективные подходы для дальнейшего улучшения непрерывного предсказания эмоций. Традиционными методами объединения предсказания базовых алгоритмов принятия решения в коллективе являются простое голосование (или голосование большинством), взвешенное голосование, взвешенное среднее, бэггинг, стекинг, бустинг и пр.

Общая схема алгоритма коллективного принятия решения состоит из следующих этапов.

1. Разбить исходную выборку на три части: обучающую, тестовую, проверочную.
2. Выбрать множество агентов для включения их в состав коллектива.
3. Обучить множество имеющихся агентов на обучающем множестве.
4. Для каждого объекта выборки из тестового множества сформировать коллективное решение из агентов с помощью среднего, взвешенного среднего или системы на нечеткой логике.
5. Вычислить ошибку сформированного коллектива на проверочном множестве.

В работе [15] разработан новый метод проектирования интеллектуальных информационных технологий, который основан на нечеткой логике и предназначен для решения задач классификации и регрессии. Нечеткий контроллер (НК) принимает решение о выборе классификационной или регрессионной модели в зависимости от близости тестового объекта к объектам из обучающей выборки и успешностью модели на ближайшем объекте.

Коллектив, спроектированный приведенным методом, апробирован при решении задачи распознавания лиц. В рамках настоящей работы предлагается применить разработанный ранее подход при решении задачи регрессии на базе данных RECOLA [16].

Модификация алгоритма коллективного принятия решения на основе нечеткой логики. Для повышения эффективности решения задачи регрессии необходимо модифицировать процедуру построения НК, а также сам алгоритм принятия решения с его помощью.

Для НК формируются три лингвистические переменные (ЛП) для входа и одна для выхода.

1. *Distance*. Близость объекта тестовой выборки к ближайшей точке из обучающей выборки. Метрика ЛП — расстояние по евклидовой метрике. Терм — множество: {близко, среднее, далеко}.
2. *Error*. Разность между выходом регрессионной модели (агента) на тестовой выборке и в ближайшей точке обучающего множества. Метрика

ЛП — расстояние по евклидовой метрике. Терм — множество: {низкая, средняя, высокая}.

3. *Weight_agent*. Вес агента (вычисляется пропорционально ошибке на обучающем множестве). Метрика ЛП — безразмерная величина в интервале от 0 до 1 (0 — ни одна из точек не была предсказана алгоритмом регрессии, 1 — были предсказаны абсолютно все точки). Сумма весов по всем агентам равна единице. Терм — множество: {низкий, высокий}.

4. *Confidence*. Степень доверия к регрессионной модели. Метрика ЛП — безразмерная величина в интервале от 0 до 1 (0 — полное отсутствие доверия к алгоритму регрессии, а 1 — абсолютное доверие). Терм — множество: {низкая, средняя, высокая}.

Выходом работы НК для каждого объекта выборки из тестового множества является степень доверия к алгоритму регрессии.

Варьируемыми параметрами пользователя являются два показателя: число ближайших точек из обучающего множества к объекту из тестовой выборки (*nPoints*), а также число агентов, композиция которых будет применяться для получения решения для объекта из тестовой выборки (*nAgent*).

При построении решающего правила для каждой тестовой точки, решение будет приниматься тем алгоритмом, степень доверия к которому больше. Эта схема работает в том случае, если принятие решения происходит по одному агенту, т. е. $nAgent = 1$.

В настоящей работе также предлагается выполнить гибридизацию методов принятия коллективного решения следующим образом. Для текущего решения с помощью системы на нечеткой логике отбирается *nAgent* лучших агентов (*nAgent* — параметр алгоритма), а из них итоговое решение принимается средним (*FLS_mean*) или взвешенным средним (*FLS_wmean*). При этом веса агентов определяются пропорционально их уверенности на данном примере.

Таким образом, система на нечеткой логике, позволяющая получить степень доверия агента, зависит от пяти параметров, переданных в НК:

$$FLS(distance, error, weight_agent, nPoint, nAgent) = \mu.$$

Функции принадлежности для всех лингвистических переменных входа и выхода имеют треугольный вид. База правил выглядит следующим образом.

IF *error* — высокая THEN *confidence* — низкая.

IF *error* — средняя AND *distance* — близко AND *weight_agent* — высокий THEN *confidence* — высокая.

IF *error* — средняя AND *distance* — среднее THEN *confidence* — средняя.

IF *error* — низкая AND *distance* — близко AND *weight_agent* — высокий THEN *confidence* — высокая.

IF *error* — низкая AND *distance* — близко AND *weight_agent* — низкий THEN *confidence* — средняя.

IF *error* — низкая AND *distance* — среднее AND *weight_agent* — высокий THEN *confidence* — высокая.

IF *distance* — далеко THEN *confidence* — низкая.

База правил является достаточно простой и интерпретируемой.

Многокритериальные алгоритмы оптимизации. Многие проблемы проектирования и поддержки принятия решений можно сформулировать как проблемы оптимизации. И в большинстве реальных задач необходимо принимать решение, основываясь не на одном критерии, или показателе качества, а на их совокупности. При решении задачи прогнозирования эмоционального поведения человека коллективными методами необходимо учитывать совокупность двух критериев: точности решения задачи и ее вычислительной сложности.

При выборе агентов в коллективе необходимо учитывать не только точность коллективного решения, но и число агентов в коллективе. Современные средства формирования классификационных и регрессионных моделей позволяют строить модели практически в автоматизированном режиме. И для одной задачи можно получить десятки эффективных агентов. Однако с ростом числа агентов в коллективе уменьшается способность к обобщению и коллектив становится склонным к переобучению. Поэтому при максимизации точности коллектива на тестовых данных необходимо минимизировать число агентов в коллективе.

В настоящее время генетические алгоритмы являются одним из наиболее эффективных средств решения задач многокритериальной оптимизации. Самыми используемыми являются алгоритмы второго поколения SPEA-II [7] и NSGA-II [6]. Эти алгоритмы отличаются относительно простыми реализацией и настройкой, а также разными принципами поиска решений. Выбранные методы основаны на идее доминирования Парето. Практика применения SPEA-II и NSGA-II и получаемые ими высокие результаты, как правило, ограничиваются двух- и трехкритериальными задачами оптимизации, поскольку при увеличении числа критериев сходимость имеет тенденцию к ухудшению точности. Причина этого видится в том, что с ростом числа критериев увеличивалась скорость заполнения популяции недоминируемыми решениями.

В настоящей работе предлагается использовать МЭА оптимизации для экономии вычислительных ресурсов и обеспечения необходимого качества работы коллективной модели при решении задачи регрессии.

При формировании коллектива необходимо выбирать, кто из агентов будет входить в состав коллектива при принятии конечного решения. Если формировать структуру состава коллектива с использованием МЭА, то необходимо закодировать ее в бинарную строку, где один агент закодирован с помощью одного бита. Длина бинарной строки постоянна, так как кодируется для максимального числа агентов, имеющих возможность включения в коллектив.

На рис. 1 приведен пример кодирования бинарной строки МЭА, длина которой равняется числу всех имеющихся агентов.

1	0	1	1	0	0	1	0	1	0
---	---	---	---	---	---	---	---	---	---

Рис. 1. Способ кодирования состава коллектива методов интеллектуального анализа данных для коллективного принятия решения (0 — агент не принимает участие в принятии решения, 1 — участвует)

В качестве критериев используется коэффициент корреляции согласованности ρ_c и число агентов, формирующих коллектив,

$$\rho_c = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\mu_x - \mu_y)^2},$$

где μ_x и μ_y — средние значения двух переменных; σ_x^2 и σ_y^2 — соответствующие дисперсии; ρ — коэффициент корреляции между двумя переменными.

Далее проводилось тестирование алгоритмов при решении задачи предсказания эмоционального состояния человека на основе базы данных RECOLA по следующим параметрам: 100 поколений, 100 индивидов.

Особенности DB-RECOLA для решения задачи прогнозирования эмоционального состояния человека по аудиоданным. Системы автоматического распознавания эмоций, основанные на контролируемом машинном обучении, требуют надежной аннотации эмоционального поведения для построения полезных моделей. В настоящей работе исследуется эффективность использования алгоритмов машинного обучения и их коллективов, способных интегрировать контекстную информацию в моделирование, для автоматического прогнозирования эмоций от нескольких (асинхронных) оценщиков в непрерывных временных областях, т. е. возбуждение и валентность. Оценка выполняется на базе дан-

ных RECOLA [16], которая содержит модуль Annotation. Этот модуль содержит аннотации (данные об аффективном поведении), выполненные шестью ассистентами (тремя мужчинами, тремя женщинами) с помощью веб-инструмента аннотации ANNEMO. Аннотации представлены для каждой аудиозаписи, записанной отдельно для каждого участника с частотой кадров 40 мс для аффективного (эмоционального) поведения.

В пространственном подходе к определению эмоций используется *n*-мерное пространство, состоящее из факторов эмоций. Любая эмоция может быть представлена как некоторая комбинация этих факторов. Для одной из таких комбинаций используются два измерения: Valence и Arousal, где первое измерение — насколько положительна или отрицательна эмоция (валентность), а второе измерение — насколько интенсивно физическое возбуждение эмоции (возбуждение) [17]. Например, Happy — это состояние с высокой валентностью, сильное возбуждение, а Stressed — это состояние с низкой валентностью и сильное возбуждение.

База данных представлена 23 субъектами, аудиозапись каждого субъекта длится 5 мин и оценивается 6 аннотаторами. На основе этого необходимо решить задачу прогнозирования аффективного (эмоционального) поведения.

Однако автоматическое распознавание эмоций от непрерывных по времени меток требует определить соответствующую длину временного окна, используемого для предсказания эмоций, которая зависит от модальности и эмоций.

В литературе нет четкого консенсуса относительно наилучшей длины временного окна, которое можно использовать для данной модальности и эмоций. В то время как предполагается, что общая продолжительность эмоции падает от 0,5 до 4 с, длина окна анализа, используемого для предсказания эмоций, может значительно варьироваться в зависимости от модальности; аудио-сигналы обычно меняются со временем быстрее, чем видеосигналы, и даже больше, чем физиологические сигналы. Выбранная последовательность преобразуется во входы с помощью скользящего окна, после чего, данные нормируются.

При увеличении размера окна существенно увеличивается размерность признакового пространства входов. В задаче RECOLA используется продолжительность временного периода 1,5 с, размер признакового пространства составляет 5200, а размерность выборки составляет более 170 000 объектов. Выбор размера временного окна сделан на основе результатов работы [18], в которой наибольшее значение точности решения задачи по аудиоданным достигалось с периодом окна в 1–2 с.

Для уменьшения признакового пространства предлагается применить метод отбора главных компонент (РСА — *principal component analysis*) с сокращением размерности до 200 компонент. После применения работы РСА исходная выборка делится на три множества: объем обучающей выборки (*learning*) 65 % , тестовой (*testing*) 25 % , а контрольной (*validation*) 15 % общего числа точек.

Результаты экспериментов. Для формирования коллективов решающих правил МЭА в задаче анализа эмоций человека по аудиоданным были взяты следующие агенты.

1. Ансамбль деревьев решений методом градиентного бустинга (GBR).

2. Алгоритм k ближайших соседей для задачи регрессии (KNR).

3. Линейная регрессия, которая строится на основе метрики L1 (LLasso).

4. Линейная регрессия, которая строится методом наименьших квадратов (LR).

5. Гребневая линейная регрессия, которая строится на основе метрики L2 (LRidge).

6–7. Искусственная нейронная сеть (многослойный перцептрон) (MLP), построенная на основе библиотеки Scikit-learn (Python). Структура сети: $200 \times 100 \times 50 \times 20$, 200×200 нейронов на соответствующих слоях, сигмоидальная активационная функция.

8–10. Искусственная нейронная сеть (многослойный перцептрон) (NNKeras), построенная на основе библиотеки Keras и TensorFlow (Python). Структура сети: $200 \times 100 \times 20$, $200 \times 100 \times 50 \times 20$, 200×200 нейронов на соответствующих слоях, сигмоидальная активационная функция.

11–12. Ансамбль деревьев принятия решений методом «случайного леса» (RFR). Число деревьев в ансамбле 10, 50. Глубина дерева 18. Число признаков, используемых одним деревом, 100.

13. Метод опорных векторов для регрессии (SVR).

Критерием эффективности работы предлагаемого подхода на основе нечеткой логики является коэффициент корреляции согласованности r_c . Число агентов, композиция которых применялась для получения решения для объекта из тестовой выборки, варьировалось от 1 до 10.

Для текущего решения с помощью системы на нечеткой логике применялось две конфигурации: итоговое решение принималось средним ($FLS+M$) и взвешенным средним ($FLS+W$).

Пример эффективности решения задачи прогнозирования метки Arousal приведен на рис. 2. Число агентов при оптимизации состава коллектива, а также без нее существенно влияет на результат. Начиная с $nAgs = 3$ точность ухудшается, соответственно параметр $nAgent$ должен быть не слишком маленьким и не слишком большим. В результате проведенных исследований (табл. 1 и 2) можно сделать выводы о том, что коллективный метод принятия решения FLS показывает большее значение точности, чем методы *mean* и *wMean*. Гибридизация метода FLS в совокупности с принятием решения на основе *mean* и *wMean* позволяет улучшить результаты FLS при эффективном выборе числа агентов, принимающих решение. Для метки Valence хорошо работают два и три агента. При этом гибридизация FLS и *wMean* работают не хуже, а в ряде случаев лучше, чем FLS и *mean*. С ростом числа агентов точность уменьшается.

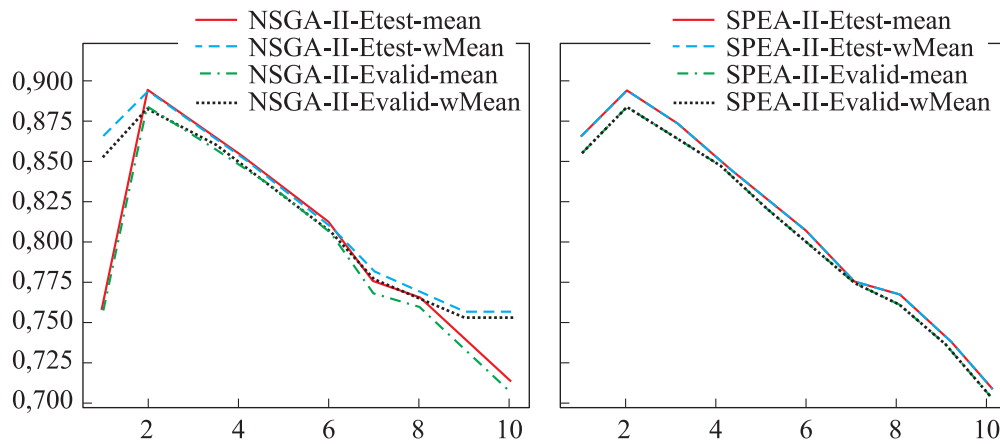


Рис. 2. Эффективность коллективов с оптимизированным составом в зависимости от выбранного параметра $nAgent$

Таблица 1

Результаты наихудшего и наилучшего агента для меток Valence и Arousal

Метка класса	Agent	Etest	Evalid
Valence	Worst Agent	GBR:0,2914	GBR:0,258
	Best Agent	NNKeras:0,796	NNKeras:0,796
Arousal	Worst Agent	GBR:0,4617	GBR:0,4617
	Best Agent	NNKeras:0,8257	NNKeras:0,8236

По результатам, приведенным в табл. 3, невозможно выбрать победителя между SPEA-II и NSGA-II. Данные алгоритмы показали сравнимую эффективность.

Таблица 2

Анализ эффективности коллективного принятия решения на основе нечеткой логики для прогнозирования метки Valence

nAg	Method	Without		NSGA-2			SPEA-2		
		Etest	Evalid	Etest	Evalid	Mask MEA	Etest	Evalid	Mask MEA
1	FLS+M	0,810	0,822	0,842	0,840	0000110101110	0,846	0,842	1001010101011
	FLS+W	0,810	0,822	0,842	0,840	0000110101110	0,846	0,842	1001010101011
2	FLS+M	0,863	0,851	0,870	0,867	0101101111110	0,873	0,872	1011111111000
	FLS+W	0,863	0,851	0,870	0,867	0101101111110	0,875	0,870	1010111111110
3	FLS+M	0,865	0,858	0,866	0,872	1001101111101	0,869	0,871	1010101111000
	FLS+W	0,865	0,859	0,866	0,872	1001101111101	0,870	0,874	0110110111010
4	FLS+M	0,840	0,832	0,841	0,843	0100101111101	0,838	0,840	0110101111110
	FLS+W	0,842	0,835	0,839	0,843	0101100111110	0,840	0,844	0111101111111
5	FLS+M	0,809	0,797	0,807	0,805	0111010111110	0,806	0,803	0101010111101
	FLS+W	0,813	0,801	0,807	0,807	1101010111110	0,812	0,812	0101001111111
6	FLS+M	0,779	0,762	0,773	0,767	0111010111110	0,773	0,768	0110100111111
	FLS+W	0,784	0,769	0,778	0,773	0111010111110	0,775	0,769	1110010111101
7	FLS+M	0,749	0,731	0,731	0,724	1110110111011	0,742	0,736	1111101111111
	FLS+W	0,755	0,738	0,740	0,735	0111010111110	0,739	0,734	0111101111110
8	FLS+M	0,718	0,699	0,699	0,691	1101101111110	0,700	0,689	0101100111111
	FLS+W	0,725	0,708	0,710	0,704	1111101111110	0,787	0,779	010000011010
9	FLS+M	0,689	0,669	0,665	0,654	1101101111110	0,665	0,656	0111101111110
	FLS+W	0,696	0,678	0,729	0,722	0000001101001	0,697	0,692	0101000110010
10	FLS+M	0,659	0,641	0,645	0,633	1111100111111	0,633	0,623	1111101111101
	FLS+W	0,668	0,651	0,733	0,720	1000010011000	0,706	0,689	1000000011110
	<i>mean</i>	0,580	0,559	0,723	0,709	1000010011000	0,706	0,688	0000001011110
	<i>wMean</i>	0,591	0,570	0,733	0,720	1000010011000	0,734	0,729	0100100110000

Таблица 3

Анализ эффективности коллективного принятия решения на основе нечеткой логики для прогнозирования метки Arousal

nAg	Method	Without		NSGA-2			SPEA-2		
		Etest	Evalid	Etest	Evalid	Mask MEA	Etest	Evalid	Mask MEA
1	FLS+M	0,848	0,843	0,759	0,757	1000000010011	0,864	0,854	0010100100001
	FLS+W	0,848	0,843	0,865	0,851	0000001101001	0,864	0,854	0010100100001
2	FLS+M	0,876	0,868	0,894	0,883	01111101110111	0,893	0,883	0011010111101
	FLS+W	0,876	0,868	0,894	0,883	0111101110111	0,893	0,883	0011010111101
3	FLS+M	0,874	0,868	0,874	0,867	0111101110111	0,875	0,866	0110100110001
	FLS+W	0,875	0,868	0,875	0,869	0111101110111	0,875	0,866	0110100110001
4	FLS+M	0,858	0,850	0,855	0,847	0010011110101	0,850	0,848	1011011110111
	FLS+W	0,859	0,852	0,855	0,851	0010011110101	0,850	0,848	1011011110111
5	FLS+M	0,838	0,833	0,832	0,831	1111010111101	0,827	0,821	1100100110111
	FLS+W	0,840	0,835	0,831	0,827	11110101111010	0,827	0,821	1100100110111
6	FLS+M	0,820	0,817	0,811	0,806	1111010111110	0,804	0,798	1101100111110
	FLS+W	0,822	0,819	0,810	0,807	1111010111100	0,804	0,798	1101100111110
7	FLS+M	0,802	0,800	0,775	0,768	1100100111111	0,774	0,774	1010111110011
	FLS+W	0,805	0,803	0,782	0,777	01011111111110	0,774	0,774	1010111110011
8	FLS+M	0,785	0,783	0,765	0,759	0100110111011	0,765	0,759	1111101110011
	FLS+W	0,788	0,786	0,770	0,763	1101101111110	0,765	0,759	1111101110011
9	FLS+M	0,768	0,766	0,741	0,735	1101101111110	0,740	0,735	1110111110011
	FLS+W	0,772	0,770	0,757	0,752	0000001101001	0,740	0,735	1110111110011
10	FLS+M	0,752	0,748	0,714	0,707	1100100111111	0,708	0,703	1100100111001
	FLS+W	0,756	0,752	0,757	0,752	0000001101001	0,708	0,703	1100100111001
	<i>mean</i>	0,679	0,676	0,766	0,756	100100111010	0,787	0,778	0000010100000
	<i>wMean</i>	0,687	0,683	0,759	0,757	1000000010011	0,782	0,775	0000100110110

Заключение. Рассмотрена задача прогнозирования эмоционального поведения человека по аудиоданным. Предложена эффективная процедура формирования коллективного решения на основе гибридизации FLS и *mean/wMean*, позволяющая получить решение точнее, чем решение лучшего агента из коллектива. Предложена процедура формирования эффективного состава коллектива с помощью многокритериального генетического алгоритма (NSGA-II и SPEA-II). В ходе исследований показано, что выбор эффективных агентов с помощью многокритериального генетического алгоритма позволяет существенно повысить точность работы и обобщающую способность коллектива. Для формирования эффективного коллектива рекомендуется использовать SPEA-II, поскольку SPEA-II в общем случае не уступает NSGA-II, а в некоторых конкретных случаях превосходит.

ЛИТЕРАТУРА

- [1] Kuncheva L. Combining pattern classifiers. Methods and algorithms. Wiley, 2004.
- [2] Hansen L.K., Salmon P. Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1990, vol. 12, no. 10, pp. 993–1001. DOI: 10.1109/34.58871
- [3] Yamaguchi T., Mackin K.J., Nunohiro E., et al. Artificial neural network ensemble-based land-cover classifiers using MODIS data. *Artif. Life Robotics*, 2009, vol. 13, no. 2, pp. 570–574. DOI: 10.1007/s10015-008-0615-4
- [4] Ridgeway G. The state of boosting. *Proc. 31st Symp. Interface*, 1999, pp. 172–181.
- [5] Breiman L. Bagging predictors. *Mach. Learn.*, 1996, vol. 24, no. 2, pp. 123–140. DOI: 10.1023/A:1018054314350
- [6] Deb K., Agrawal S., Pratap A., et al. A fast elitist non-dominated sorting genetic algorithm for Multi-objective optimization: NSGA-II. *Parallel Problem Solving from Nature PPSN VI*. Springer, 2000, pp. 849–858. DOI: 10.1007/3-540-45356-3_83
- [7] Zitzler E., Laumanns M., Thiele L. SPEA2: improving the strength Pareto evolutionary algorithm 2. TIK Report 103. Computer Engineering and Networks Laboratory, Swiss Federal Institute of Technology (ETH). Zurich, Switzerland, 2001.
- [8] Gunes H., Pantic M. Automatic, dimensional and continuous emotion recognition. *IJSE*, 2010, vol. 1, no. 1, pp. 68–99. DOI: 10.4018/jse.2010101605
- [9] Drucker H., Burges C.J., Kaufman L., et al. Support vector regression machines. *Adv. Neural Inf. Process. Syst.*, 1997, vol. 9, pp. 155–161.
- [10] Tipping M.E. Sparse Bayesian learning and the relevance vector machine, *JMLR*, 2001, vol. 1, pp. 211–244.
- [11] Kwok T.-Y., Yeung D.-Y. Constructive algorithms for structure learning in feed-forward neural networks for regression problems. *IEEE Trans. Neural Netw.*, 1997, vol. 8, no. 3, pp. 630–645. DOI: 10.1109/72.572102

- [12] Williams R.J., Zipser D. A learning algorithm for continually running fully recurrent neural networks. *Neural Comput.*, 1989, vol. 1, no. 2, pp. 270–280. DOI: 10.1162/neco.1989.1.2.270
- [13] Tian L., Moore J.D., Lai C. Emotion recognition in spontaneous and acted dialogues. *Proc. ACII*, 2015, pp. 698–704. DOI: 10.1109/ACII.2015.7344645
- [14] Nicolaou M.A., Gunes H., Pantic M. Continuous prediction of spontaneous affect from multiple cues and modalities in valence-arousal space. *IEEE Trans. Affect. Comput.*, 2011, vol. 2, no. 2, pp. 92–105. DOI: 10.1109/T-AFFC.2011.9
- [15] Polyakova A., Lipinskiy L. A study of fuzzy logic ensemble system performance on face recognition problem. *IOP Conf. Ser.: Mater. Sci. Eng.*, 2017, vol. 173, no. 1, art. 012013. DOI: 10.1088/1757-899X/173/1/012013
- [16] Ringeval F., Sonderegger A., Sauer J., et al. Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions. *10th IEEE Int. Conf. Workshops on Automatic Face and Gesture Recognition (FG)*, 2013, pp. 1–8. DOI: 10.1109/FG.2013.6553805
- [17] Russell J.A. A circumplex model of affect. *J. Pers. Soc. Psychol.*, 1980, vol. 39, no 6, pp. 1161–1178. DOI: 10.1037/h0077714
- [18] Ringeval F., Eyben F., Kroupi E., et al. Prediction of asynchronous dimensional emotion ratings from audio-visual and physiological data. *Pattern Recogni. Lett.*, 2015, vol. 66, pp. 22–30. DOI: 10.1016/j.patrec.2014.11.007

Полякова Анастасия Сергеевна — аспирант кафедры «Системный анализ и исследование операций» Сибирского государственного университета науки и технологий имени академика М.Ф. Решетнева (Российская Федерация, 660037, Красноярский край, г. Красноярск, пр-т им. газеты Красноярский рабочий, д. 31).

Липинский Леонид Витальевич — канд. техн. наук, доцент кафедры «Системный анализ и исследование операций» Сибирского государственного университета науки и технологий имени академика М.Ф. Решетнева (Российская Федерация, 660037, Красноярский край, г. Красноярск, пр-т им. газеты Красноярский рабочий, д. 31).

Просьба ссылаться на эту статью следующим образом:

Полякова А.С., Липинский Л.В. Формирование коллектива решающих правил многокритериальным эволюционным алгоритмом в задаче анализа эмоций человека по аудиоданным. *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*, 2019, № 4, с. 45–61. DOI: 10.18698/0236-3933-2019-4-45-61

DECISION RULE ENSEMBLE FORMATION VIA A MULTICRITERIA EVOLUTIONARY ALGORITHM FOR THE PROBLEM OF HUMAN EMOTION ANALYSIS IN AUDIO DATA

A.S. Polyakova

polyakova_nasty@mail.ru

L.V. Lipinskiy

lipinskiyl@mail.ru

**Reshetnev Siberian State University of Science and Technology,
Krasnoyarsk, Russian Federation**

Abstract

One of the most important problems at the current stage of social informatisation is development of human-machine interaction systems, including automated human emotion recognition systems. It is possible to describe human emotions using a combination of two parameters: Valence, which represents how attractive an emotion is (referring to positive and negative emotions), and Arousal, denoting the strength of the emotion (that is, degree of agitation). These parameters are real numbers. We propose to employ ensemble learning methods to improve prediction accuracy. We evaluate the accuracy of an ensemble decision via its congruence coefficient. We used a multicriteria evolutionary algorithm to select agents (algorithms) for the ensemble. Employing a multicriteria evolutionary algorithm made it possible to automate the ensemble formation process, which enabled us to save time and physical resources. Ensemble formation depended on two criteria: maximising accuracy and minimising the number of agents in the ensemble. We used the following ensemble decision-making methods: majority voting, weighted average, weighted average in proportion to the agent trust, and a fuzzy logic system. We present a modification to the fuzzy logic system that improves solution efficiency for the data mining problem. We analysed and investigated how efficient a multicriteria evolutionary algorithm is when solving the problem of predicting emotional behaviour in humans. Our experiments showed that using a multicriteria evolutionary algorithm to automate ensemble formation improves the solution accuracy

Keywords

Ensemble, evolutionary computation, multicriteria optimization algorithms, data mining algorithms, fuzzy logic systems, regression problem

Received 15.01.2019

© Author(s), 2019

The study was supported by the Ministry of Education and Science of the Russian Federation as part of basic state funding of project no. 2.1680.2017/PCh

REFERENCES

- [1] Kuncheva L. Combining pattern classifiers. Methods and algorithms. Wiley, 2004.
- [2] Hansen L.K., Salmon P. Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1990, vol. 12, no. 10, pp. 993–1001. DOI: 10.1109/34.58871
- [3] Yamaguchi T., Mackin K.J., Nunohiro E., et al. Artificial neural network ensemble-based land-cover classifiers using MODIS data. *Artif. Life Robotics*, 2009, vol. 13, no. 2, pp. 570–574. DOI: 10.1007/s10015-008-0615-4
- [4] Ridgeway G. The state of boosting. *Proc. 31st Symp. Interface*, 1999, pp. 172–181.
- [5] Breiman L. Bagging predictors. *Mach. Learn.*, 1996, vol. 24, no. 2, pp. 123–140. DOI: 10.1023/A:1018054314350
- [6] Deb K., Agrawal S., Pratap A., et al. A fast elitist non-dominated sorting genetic algorithm for Multi-objective optimization: NSGA-II. *Parallel Problem Solving from Nature PPSN VI*. Springer, 2000, pp. 849–858. DOI: 10.1007/3-540-45356-3_83
- [7] Zitzler E., Laumanns M., Thiele L. SPEA2: improving the strength Pareto evolutionary algorithm 2. TIK Report 103. Computer Engineering and Networks Laboratory, Swiss Federal Institute of Technology (ETH). Zurich, Switzerland, 2001
- [8] Gunes H., Pantic M. Automatic, dimensional and continuous emotion recognition. *IJSE*, 2010, vol. 1, no. 1, pp. 68–99. DOI: 10.4018/jse.2010101605
- [9] Drucker H., Burges C.J., Kaufman L., et al. Support vector regression machines. *Adv. Neural Inf. Process. Syst.*, 1997, vol. 9, pp. 155–161.
- [10] Tipping M.E. Sparse Bayesian learning and the relevance vector machine, *JMLR*, 2001, vol. 1, pp. 211–244.
- [11] Kwok T.-Y., Yeung D.-Y. Constructive algorithms for structure learning in feed-forward neural networks for regression problems. *IEEE Trans. Neural Netw.*, 1997, vol. 8, no. 3, pp. 630–645. DOI: 10.1109/72.572102
- [12] Williams R.J., Zipser D. A learning algorithm for continually running fully recurrent neural networks. *Neural Comput.*, 1989, vol. 1, no. 2, pp. 270–280. DOI: 10.1162/neco.1989.1.2.270
- [13] Tian L., Moore J.D., Lai C. Emotion recognition in spontaneous and acted dialogues. *Proc. ACII*, 2015, pp. 698–704. DOI: 10.1109/ACII.2015.7344645
- [14] Nicolaou M.A., Gunes H., Pantic M. Continuous prediction of spontaneous affect from multiple cues and modalities in valence-arousal space. *IEEE Trans. Affect. Comput.*, 2011, vol. 2, no. 2, pp. 92–105. DOI: 10.1109/T-AFFC.2011.9
- [15] Polyakova A., Lipinskiy L. A study of fuzzy logic ensemble system performance on face recognition problem. *IOP Conf. Ser.: Mater. Sci. Eng.*, 2017, vol. 173, no. 1, art. 012013. DOI: 10.1088/1757-899X/173/1/012013
- [16] Ringeval F., Sonderegger A., Sauer J., et al. Introducing the RECOLA multimodal corpus of remote collaborative and affective interactions. *10th IEEE Int. Conf. Workshops on Automatic Face and Gesture Recognition (FG)*, 2013, pp. 1–8. DOI: 10.1109/FG.2013.6553805

[17] Russell J.A. A circumplex model of affect. *J. Pers. Soc. Psychol.*, 1980, vol. 39, no. 6, pp. 1161–1178. DOI: 10.1037/h0077714

[18] Ringeval F., Eyben F., Kroupi E., et al. Prediction of asynchronous dimensional emotion ratings from audio-visual and physiological data. *Pattern Recogni. Lett.*, 2015, vol. 66, pp. 22–30. DOI: 10.1016/j.patrec.2014.11.007

Polyakova A.S. — Post-Graduate Student, Department of Systems Analysis and Operations Research, Reshetnev Siberian State University of Science and Technology (Imeni gazety Krasnoyarskiy rabochiy prospekt 31, Krasnoyarsk, 660037 Russian Federation).

Lipinskiy L.V. — Cand. Sc. (Eng.), Assoc. Professor, Department of Systems Analysis and Operations Research, Reshetnev Siberian State University of Science and Technology (Imeni gazety Krasnoyarskiy rabochiy prospekt 31, Krasnoyarsk, 660037 Russian Federation).

Please cite this article in English as:

Polyakova A.S., Lipinskiy L.V. Decision rule ensemble formation via a multicriteria evolutionary algorithm for the problem of human emotion analysis in audio data. *Herald of the Bauman Moscow State Technical University, Series Instrument Engineering*, 2019, no. 4, pp. 45–61 (in Russ.). DOI: 10.18698/0236-3933-2019-4-45-61