Дата подання: 30.09.2022

Дата прийняття:

**УДК 621.391**

**Автоматическое построение нечёткой MISO системы при помощи матричного представления правил и анализ корректности полученной системы**

**Д. И. Егошкин, Н. А. Гук**

**АВТОМАТИЧНА ПОБУДОВА НЕЧІТКОЇ MISO СИСТЕМИ ЗА ДОПОМОГОЮ МАТРИЧНОГО ПРЕДСТАВЛЕННЯ ПРАВИЛ ТА АНАЛІЗ КОРЕКТНОСТІ ОТРИМАНОЇ СИСТЕМИ**

**Д. І. Єгошкін, Н. А. Гук**

**AUTOMATIC CONSTRUCTION OF A FUZZY MISO SYSTEM USING A MATRIX REPRESENTATION OF THE RULES AND CORRECTNESS ANALYSIS OF THE RESULTING SYSTEM**

**Danylo Yehoshkin, Natalia Guk**

*Разработан поход к автоматическому созданию и последующей проверке правильности нечеткой модели логического вывода для задачи классификации. База знаний экспертной системы формируется с использованием объектов обучающей выборки на основании лингвистических переменных и терм-множеств. Применяется продукционная модель представления знаний, объединяющая модели Мамдани и Такаги-Сугено-Канг для формировании нечёткой базы знаний. Предполагается, что левые части продукционных правил описывают сочетания признаков объектов, а правые части соответствуют классам. Для автоматизации процесса построения правил, используется матричное представление антецедентов правил. Консеквенты правил представляются в виде вектора столбца. Для автоматического построения матрицы антецедентов предлагается использовать декартово произведение. Формирование вектора консеквентов осуществляется автоматически с использованием данных о предметной области и процедуры обучения. Анализа правильности базы знаний на корректность осуществляется при помощи критериев полноты, минимальности, связности и непротиворечивости. Для доказательства полноты базы правил применяется логика Хоара и метод резолюций. После чего используется программное обеспечение Simplify* *для автоматической проверки базы правил на основе метода резолюций. Simplify* *осуществляет* *доказательство истинности заданных предикатов при помощи логики первого порядка и введённых аксиом и предикатов. Полученная экспертная система применяется для решения задачи классификации. Проводится оценка качества классификации при помощи метрик: accuracy, precision, recall, f1-score. В работе проведен анализ результатов классификации и качества полученной базы знаний экспертной системы. Рассмотрена зависимость результата классификации, от размера обучающей выборки.*

*Ключевые слова: экспертная система, нечеткая логика, классификация, корректность, логика Хоара, метод резолюций.*

**1. Вступление**

Известно, что большинство современных подходов к машинному обучению оказали большое влияние на многие отрасли. Развитие технологий позволило широко применять системы искусственного интеллекта для поддержки принятия решений в таких областях знаний как экономика, промышленность, медицина, наука, торговля, строительство, транспорт. Информация о предметных областях представляет собой большие массивы данных, которые необходимо обработать для принятия правильных решений. Поэтому для решения широкого класса практических задач применяются такие методы искусственного интеллекта как нечеткая логика, нейронные сети, генетические алгоритмы и другие. Это позволяет избежать применения точных математических моделей, основанных на аппарате математических уравнений и классической теории принятия решений. С помощью искусственного интеллекта создаются системы для имитации деятельности экспертов в различных областях. Задачи управления, идентификации, моделирования сложных физических явлений, классификации, распознавания образов успешно решаются с использованием элементов нечеткой логикии [1]. Применение нечеткой логики обусловлено тем, что она на интуитивном уровне соответствует процессу человеческого рассуждения в условиях нечеткости и неполноты условий задачи. После формирования продукционной системы правил необходимо проверить данную систему на корректность. Система правил может быть неполной, несвязной и противоречивой, поэтому актуальным является разработка методов, которые позволяют осуществлять процедуру проверки и доказательства правильности полученной системы правил. Полученные экспертные системы применяются для решения практических задач: управления, принятия решений, классификации, кластеризации, data mining, big data analytics.

**2. Анализ литературных данных и постановка проблемы**

Автоматическое формирование продукционной базы правил в экспертных системах и доказательство ее корректности является актуальной задачей, поскольку позволяет обеспечить простоту разработки таких систем и высокое качество логического вывода. В литературе для решения подобных проблем развиваются разные подходы.

Для анализа достоверности базы знаний нечеткого логического вывода в работе [2] рассматривается представление системы правил в виде метаграфа. При этом нечеткая база знаний представляется с использованием модели Мамдани. После процедуры верификации база правил должна удовлетворять следующим условиям: неизбыточности, лингвистической непротиворечивости, отсутствия зацикливания и лингвистической полноты. Предложено проводить статическую верификацию баз нечетких знаний на основе структуры метаграфа путем нахождения зацикливания в структуре графа.

В [3] рассматривается вопрос автоматизации проверки корректности базы знаний продукционных правил. Предлагается квантовая модель кодирования символов многозначного алфавита при помощи кубического исчисления. В качестве алгоритма вывода используется алгоритм Мамдани. Разрабатывается компьютерная программа, которая позволяет выполнить формальную проверку базы продукционных правил на корректность при помощи метода резолюций. Теоретико-множественные операции над символами алфавита сводятся к соответствующим поразрядным логическим операциям над их кубитами. К недостаткам данного подхода можно отнести следующее: входные лингвистические переменные должны иметь одинаковое количество термов, а термы разных входных лингвистических переменных должны иметь одинаковые диапазоны значений. Кроме того, продукционные правила должны быть полностью определены, т.е. любому сочетанию значений термов входных лингвистических переменных соответствует определенное значение термов выходной лингвистической переменной.

Также для представления базы знаний применяются ациклические графы. В работе [4] для аккумуляции больших наборов знаний в реальном времени применяется алгоритм построения ориентированного ациклического графа. Последующая трансляция графа осуществляется при помощи специализированного языка TLC (Target Language Compiler). В работе предлагается проводить проверку базы знаний на корректность на этапе трансляции. Применяются как классические методы с использованием дизъюнктивных и конъюнктивных нормальных форм, так и с использованием нормальной формы отрицания NNF (Negation Normal Form). Это позволяет выводить пользователю контрпримеры, если база знаний является противоречивой или не полной.

В случае если необходимо представить данные в табличном виде для применения матричных и векторных операций обработки, используется таблица принятия решений (ТПР) [5,6]. Большим преимуществом ТПР является наглядность и удобное представление данных для эсперта в процессе анализа и заполнения базы правил. Также достоинством является то, что основными операциями продукционного цикла выступают логические (векторные и матричные) операции, позволяющие получить высокую скорость обработки. Это позволяет весьма эффективно использовать табличную модель представления знаний. Для проверки базы знаний используется дизъюнктивная нормальная форма, удобна для автоматического доказательства теорем. Процесс доказательства основывается на логике высказываний и логике предикатов.

Для генерации базы правил типа Мамдани в работе [7] предлагается определение оптимальных консеквентов на основе использования мультиагентных оптимизационных алгоритмов, в частности муравьиных. Данный подход позволяет эффективно генерировать базы правил в следующих случаях: 1) при недостаточном объеме исходной информации; 2) при достаточно большом количестве правил, для которых составление базы нечетких правил на основе знаний экспертов не всегда эффективно; 3) при различном уровне квалификации экспертов.

Указанные подходы существенно повышают эффективность использования интеллектуальных систем. Позволяют формировать и проводить верификацию базы продукционных правил, однако: 1) накладывают ограничения на количество термов лингвистических переменных; 2) требуют наличие эксперта для анализа работы верификации; 3) накладывают ограничения на количество логических выводов.

В работе рассматривается задача классификации. Для решения поставленной задачи применяется подход, основанный на нечёткой логике.

Информация о предметной области представляется в виде продукционной модели в предположении, что левые части правил продукций описывают сочетания признаков объектов, а правые соответствуют классам.

При формировании левых частей продукций используются признаки объектов из некоторого конечного (квазиконечного, если допустимо пополнение модели) множества, конъюнкция истинных значений которых и определяет условия применимости продукции. В правой части указываются классы из некоторого конечного допустимого множества классов объектов.

Для обеспечения достоверности логического вывода на основе сформулированной базы знаний необходимо, чтобы база знаний обладала свойствами полноты, минимальности (не избыточности), непротиворечивости и связанности. Разрабатывается подход, который обеспечивает проверку этих свойств.

**3. Цель и задачи исследования**

Целью исследования является разработка подхода к автоматической генерации нечетких продукционных правил на основе обучающей выборки для решения задачи классификации с последующей проверкой корректности построенной модели.

Для достижения цели были поставлены следующие задачи:

– создать продукционную модель представления знаний об объектах предметной области, объединяющую модели Мамдани и Такаги-Сугено-Канг, на основе результатов наблюдения за объектами предметной области.

– разработать алгоритм автоматического формирования и выбрать способ представления базы правил.

– разработать подход для проверки корректности базы правил с использованием логики Хоара и метода резолюций. Применить программное обеспечение Simplify для автоматического доказательства непротиворечивости;

– выбрать метрики для оценки качества полученной нечёткой базы правил;

– применить разработанный подход для проверки корректности базы правил для решения задачи классификации, выполнить анализ результатов вычислительного эксперимента.

**4. Материалы и методы исследования**

Рассматривается задача автоматического формирования нечетких продукционных правил на основе обучающей выборки для решения задачи классификации. Формулируется задача следующим способом: необходимо на основе обучающей выборки построить нечеткие продукционные правила для MISO системы. При этом формирование правил строится на основе лингвистических переменных и их терм-множеств.

Для построения антецедентов правил использованы признаки объектов из конечного терм-множества, конъюнкция истинных значений которых определяет условия применимости продукции. Антецеденты правил представляются в виде матрицы, это необходимо для автоматического формирования матрицы при помощи декартового произведения. Для консеквентов указываются классы из некоторого конечного допустимого множества классов объектов. Консеквенты правил представляются в виде вектора столбца, значения элементов вектора зависят от матрицы антецедентов и объектов обучающей выборки. Формирование вектора консеквентов осуществляется при помощи процедуры обучения базы знаний.

Для проверки полученной базы знаний на корректность, используются критерии: полноты, минимальности, связности и непротиворечивости. Для доказательства полноты применяется логика Хоара, метод резолюций и приложение Simplify, которое осуществляет автоматическую проверку системы на непротиворечивость с использованием логики первого порядка.

Разработанный подход применяется к решению задачи классификации, оценка качества полученной нечёткой базы правил проводится при помощи метрик accuracy, precision, recall, f1-score.

**5. Результаты разработки методики анализа корректности продукционной модели**

**5. 1. Продукционная модель представления знаний и метод логического вывода**

Объекты предметной области в задаче классификации описываются системой признаков  для каждого объекта *xi ∈ X*, признаки соответствуют лингвистическим переменным *k1, k2,…,kL*, при этом каждая лингвистическая переменная содержит терм множество *Alt*, где *l* номер признака , а ; *Tl* – количество термов для признака *kl*; Выходная переменная  соответствует классу, которому принадлежит объект предметной области. Для решения задачи классификации построим продукционную модель представления знаний, объединяющую модели Мамдани и Такаги-Сугено-Канг [8].

Правила *Пp* типа MISO имеют вид:

: IF (1)



где *k1, k2,…,kL* – входные лингвистические переменные;  – переменная, принимающая нечеткие значения *Alt* для правила *p*; *Alt* – нечеткое терм-множество лингвистической переменной *kl* для правила *p*; *t* – номер терма для *kl*, ; *Tl* – количество термов для признака *kl* ; *y* – выходная переменная; *p* – номер правила в базе правил, ; *Р* – общее количество правил *П*; *cp* – переменная, принимающая нечеткие значения *Cm* для правила *p*; *Cm* – метка нечеткого класса, к которому принадлежит объект *xi ∈ X*; *m* – номер класса *С*, ; *M* – количество классов *C*.

Вещественная, неотрицательная, нормированная функция *f(k1,…,kL,p)* непрерывна на отрезке [a,b] и используется в качестве весового коэффициента для консеквента правила. Она определяет степень принадлежности выходной переменной *y* к терм-множеству *Cm* и вычисляется следующим образом:

 (2)

Процедура фазизфикаци определяется следующим образом:

 (3)

где  - функция принадлежности четкого значения входной переменной *kl* нечеткому терму *apl*.

Для автоматизации процесса построения правил, удобно представить антецеденты правил в виде матрицы A:

 (4)

где *apl* – элемент матрицы, принимающий значение  – переменной из системы правил, и равный нечеткому значению *Alt* для правила *p*; P количество правил; L – размерность вектора признаков объекта *xi ∈ X*.

Консеквенты правил будем представлять в виде вектора столбца :

 (5)

Далее, нечеткое множество  классов, что отвечают выходной переменной , определяется в виде:

 Таким образом, множество  - будет состоять из упорядоченных пар *<fm* *|Cm >*, таких что:



где пара *<fm |Cm >* - определяет степень принадлежности *fm* объекта *yi* классу *Cm*.

Для дефазификации  применяется алгоритм Мамданди с центроидным методом. Нечеткое множество  классов, соответствующее входному объекту *X* определяется в виде:



где операции импликации imp и агрегирования agg реализуются путем нахождения min и max соответственно.



Четкое значение логического вывода C\*определяется в результате дефазификации нечеткого множества  по методу центра тяжести.

**5. 2. Алгоритм автоматического формирования базы правил**

Рассмотрим процесс автоматического формирования модели представления знаний в виде матрицы антецедентов A и вектора консеквентов .

Для каждой лингвистической переменной *kl*, существует конечное (квазиконечного, если допустимо пополнение модели) нечеткое терм-множество *Al=(Al1, Al2, Al3,…, Alt)* – определяющее конечный алфавит для описания состояний или множества состояний лингвистической переменной.

Количество нечетких классов *Cm*, к которым принадлежит объект *xi ∈ X* также ограничено. Таким образом, элементы *Alt* и *Cm* определяют алфавит продукционной модели.

Формально матрица антецедентов *A* задаётся упорядоченным набором *<A11,…,Alt>,* где каждая вектор-строка матрицы *A* представляет собой множество условий продукционного правила. Вектор консеквентов  состоит из классов *Cm* и имеет размерность равную количеству правил, задает соответствие между продукционным правилом в матрице *A* и нечетким классом *Cm*.

Множество продукционных правил *П* можно представить в виде: *П*p=<*ap,cp>*, где *ap* – вектор-строка матрицы *A.* Терм-множество *Al=(Al1, Al2,* *Al3,…,Alt)* – упорядочено относительно семантического порядка и границ термов. Для построения матрица *A* предлагается использовать декартово произведение:



После чего, данное множество *Axl*, можно представить в виде следующей матрицы *A*:

 (6)

Количество правил P определяется из свойств декартового произведения и равно произведению количества элементов множеств-сомножителей:

 (7)

Формирование вектора консеквентов осуществляется при помощи процедуры обучения базы знаний. Каждый объект из обучающей выборки *XTrain* проходит процедуру фазификации *fuzz(xiTrain)*, после чего вектор  определяется следующим образом:

 (8)

где *cord() –* функция мощности множества.

Вектор  должен содержать в себе все классы *Cm*, в противном случае на вход подана неполная или неверная обучающая выборка и необходимо переобучить систему на другом наборе данных.

**5. 3. Анализ корректности базы знаний**

База знаний должна удовлетворять единым формальным требованиям корректности, а именно, быть полной, минимальной, связной и непротиворечивой [3].

Под полнотой следует понимать, что любому входящему сочетанию значений термов входных лингвистических переменных соответствует определенное правило в базе правил. Минимальная база правил - это база из которой нельзя удалить ни одного из продукционных правил, не нарушив тем самым её полноту. База знаний является непротиворечивой (согласованной), если она не содержит несовместные правила - правила с одинаковыми лингвистическими условиями, но различными заключениями. Минимальность и непротиворечивость формируемой базы знаний обеспечивается свойствами декартового произведения. За связность отвечает вектор построенный в процессе обучения на основе обучающей выборки (8). Каждая строка матрицы (4) отличаются только одним из *l* подусловий, так как в начале формирования матрицы A терм-множество *Al=(Al1, Al2,* *Al3,…,Alt)* упорядочено относительно семантического порядка и границ термов.

Для автоматического доказательства полноты базы правил применяется логика Хоара [9]. Определим тройку Хоара: {Q} S {R}, где Q – предикат предусловия, S – набор команд, R – предикат постусловия. Под набором команд S будем понимать множество продукционных правил *П*p=<*ap,cp>*:

S: *<A,>.*

Предусловием в данном предикате описываются все возможные начальные значения, которые могут принимать лингвистические переменные, для замыкания множества принадлежности входных значений:



Предикат Q = T – истина, так как  принадлежит хотя бы одному *Alt*.

Предикат постусловия описывает ожидаемый результат выполнения S:

R: 

Если предусловие Q выполняется, команда S делает верным постусловие R [10].

Данное утверждение, доказывается с использованием предиката слабейшего предусловия WP(S,R) [11]:

 (9)

Предикат WP выполняет подстановку  из S в R.

 (10)

Для доказательства истинности предиката (10) воспользуемся методом резолюций:











Так как  - принимают значения из *Cm*, после подстановки S, получаем:



Далее правило резолюции применяется ко второму и третьему дизъюнктам:



Учитывая (7), получим:



Применим закон дистрибутивности ко второму дизъюнкту и осуществим подстановку S:





В результате применения правила резолюций к первому и второму дизъюнктам получим: {\_}. Приведенное доказательство показывает полноту постусловий на основании предикат WP.

Для автоматической проверки базы правил будем использовать программное обеспечение Simplify [12], в котором реализован метод резолюций для доказательства истинности заданных предикатов на основании логики первого порядка. Формальное исчисление допускает высказывания относительно переменных, фиксированных функций и предикатов, что в свою очередь расширяет логику высказываний. В отличие от других подобных средств Simplify никогда не образует бесконечный цикл в процессе доказательства и считает, что предикаты, доказательство истинности которых невозможно, являются ложными [13].

Сформулируем тройку Хоара: ; S: *<A,>;* R: 

Далее, для доказательства правильности сформулируем предикаты с использованием директив языка Simplify и проведём автоматическое доказательство.

Аксиома Simplify для первого конъюнкта предусловия Q, будет иметь вид:

(BG\_PUSH (FORALL (k\_i ai1 ai2 ai3 ... ait)

(OR (EQ k\_i ai1) (EQ k\_i ai2) (EQ k\_i ai3) ... (EQ k\_i ait))))

Аксиому для второго конъюнкта предусловия Q представим в виде:

(BG\_PUSH (FORALL (result c1 c2 c3 ... cm)

(OR (EQ result c1) (EQ result c2) (EQ result c3) ... (EQ result cm))))

После чего необходимо перевести Q и R на язык Simplify.

С использованием введённых аксиом и предикатов Q, S, R – организовано автоматическое доказательство правильности базы знаний.

**5. 4. Выбор метрик для оценки качества базы правил**

Поскольку для решения задачи используется экспертная система с нечёткой базой правил. То полученный результат классификации является приближенными, поэтому необходимо иметь возможность оценить качество полученного результата. Для оценки качества результата классификации, проведенного с использованием нечеткой базы правил, в работе используются следующие метрики:

; ; ;

,

где TP- True Positive; FP - False Positive; FN - False Negative; TN – True Negative.

**5. 5. Анализ результатов**

Предложенный подход был применён для задачи классификации видовой популяции арктических пингвинов [14]. Набор данных содержит данные о 344 особях трех видов Chinstrap, Adelie, Gentoo. Признаки особей: bill\_length\_mm – длинна клюва; bill\_depth\_mm – глубина клюва; flipper\_length\_mm — длина плавника; body\_mass\_g — масса тела; Для проведения классификации рис. 3 и рис. 4 были сформированы обучающая и тестовая выборки в размере 300/44 (обучающая/тестовая). На рис. 1 представлена диаграмма предельного распределения.

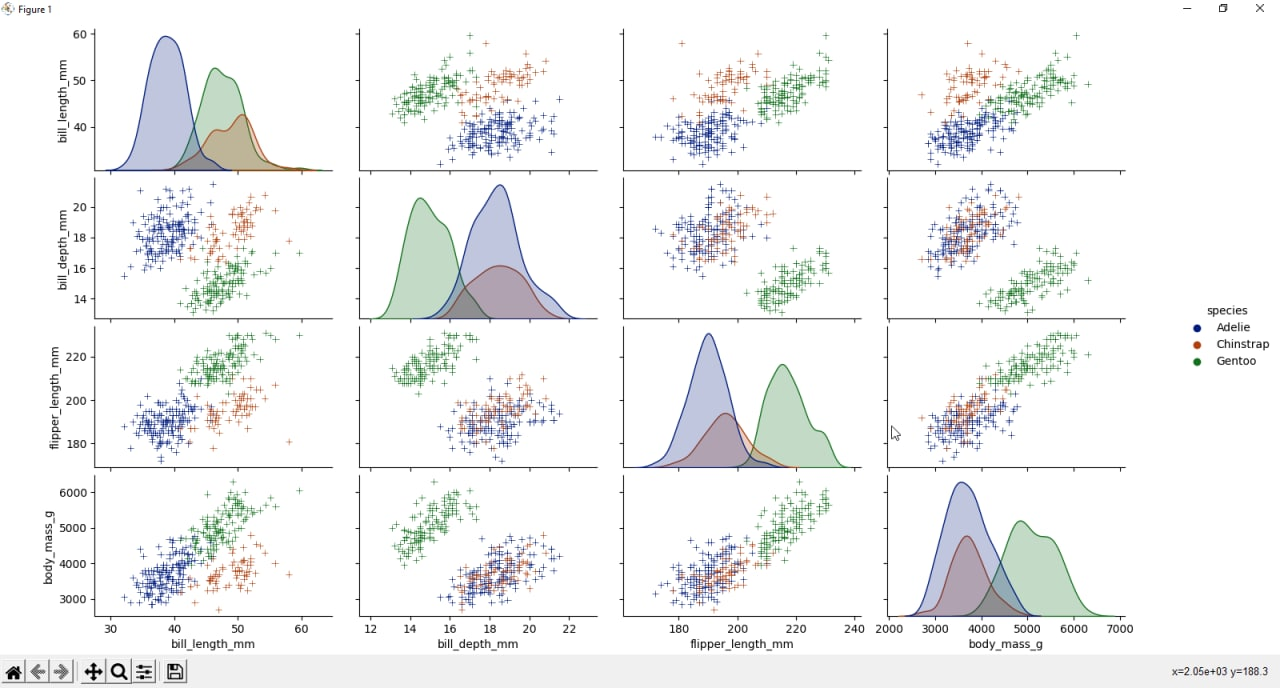


Рис. 1. Диаграмма предельного распределения классов: Chinstrap, Adelie, Gentoo

На рис. 2 приведен результат проверки базы знаний на полноту при помощи предиката слабейшего постусловия и приложения Simplify.

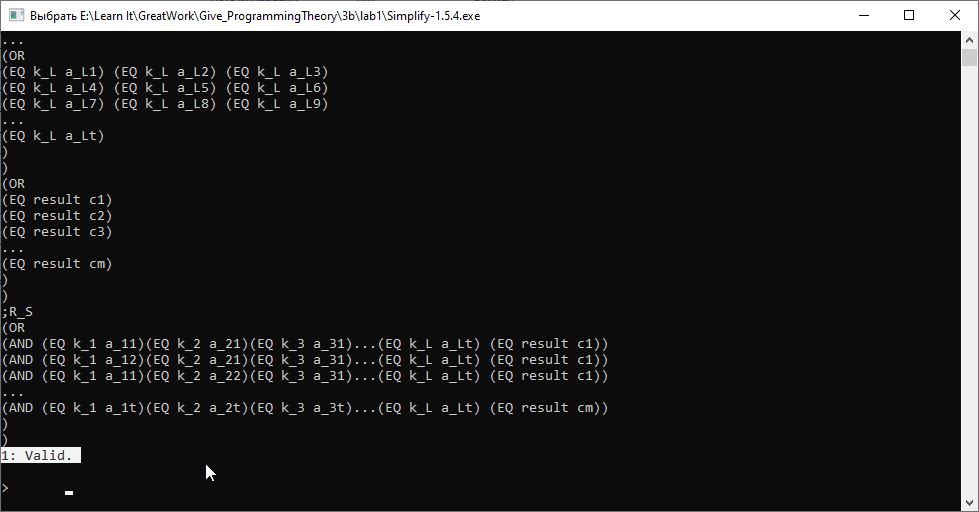


Рис. 2. Результат проверки полноты, «Valid» – истина.

На рис. 3 показаны результаты классификации до обучения нечеткой базы знаний:

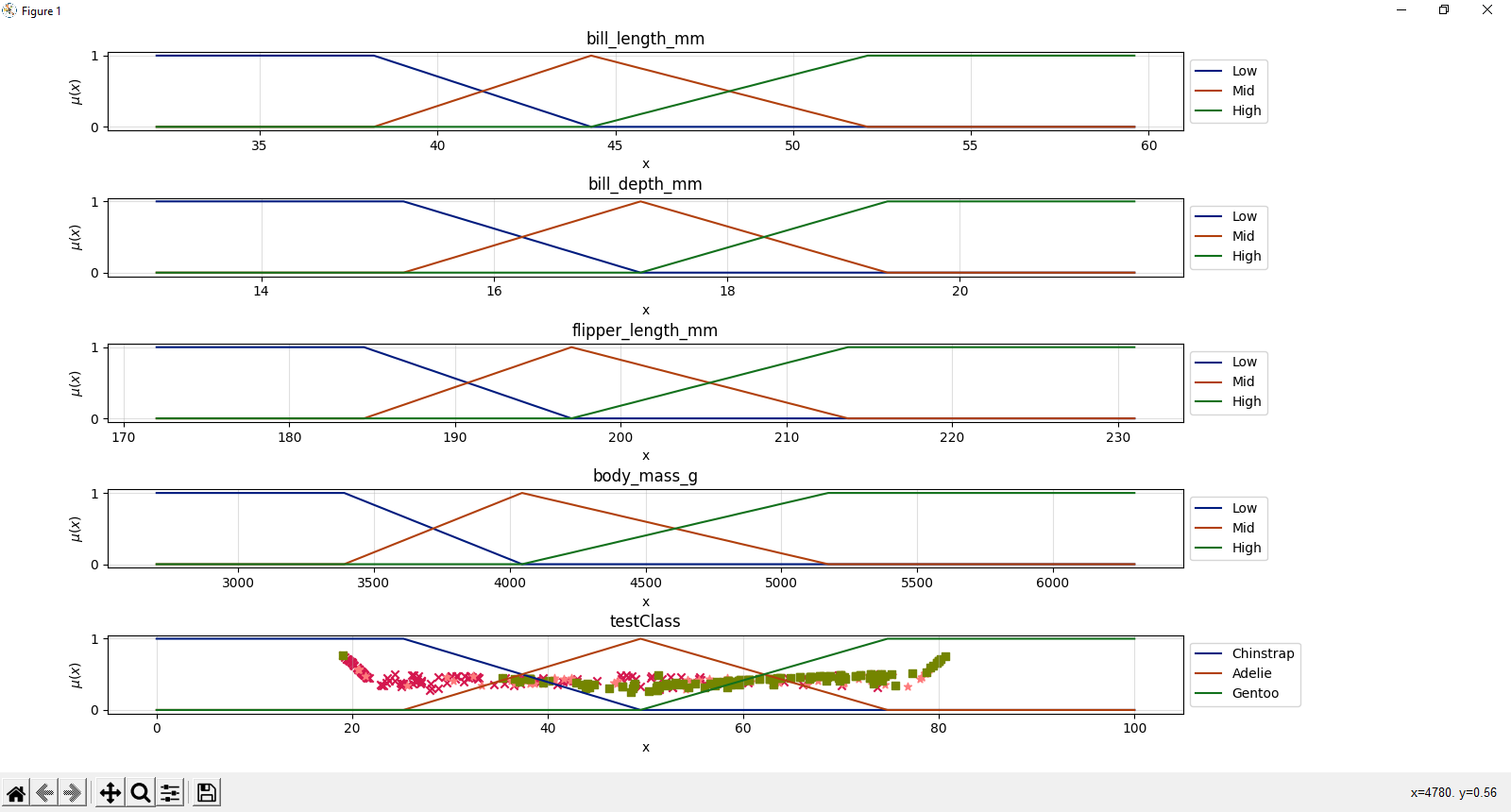


Рис. 3. До обучения: Графики 4-х функций принадлежности и результата

На рис. 4 Приведены результаты классификации после обучения:

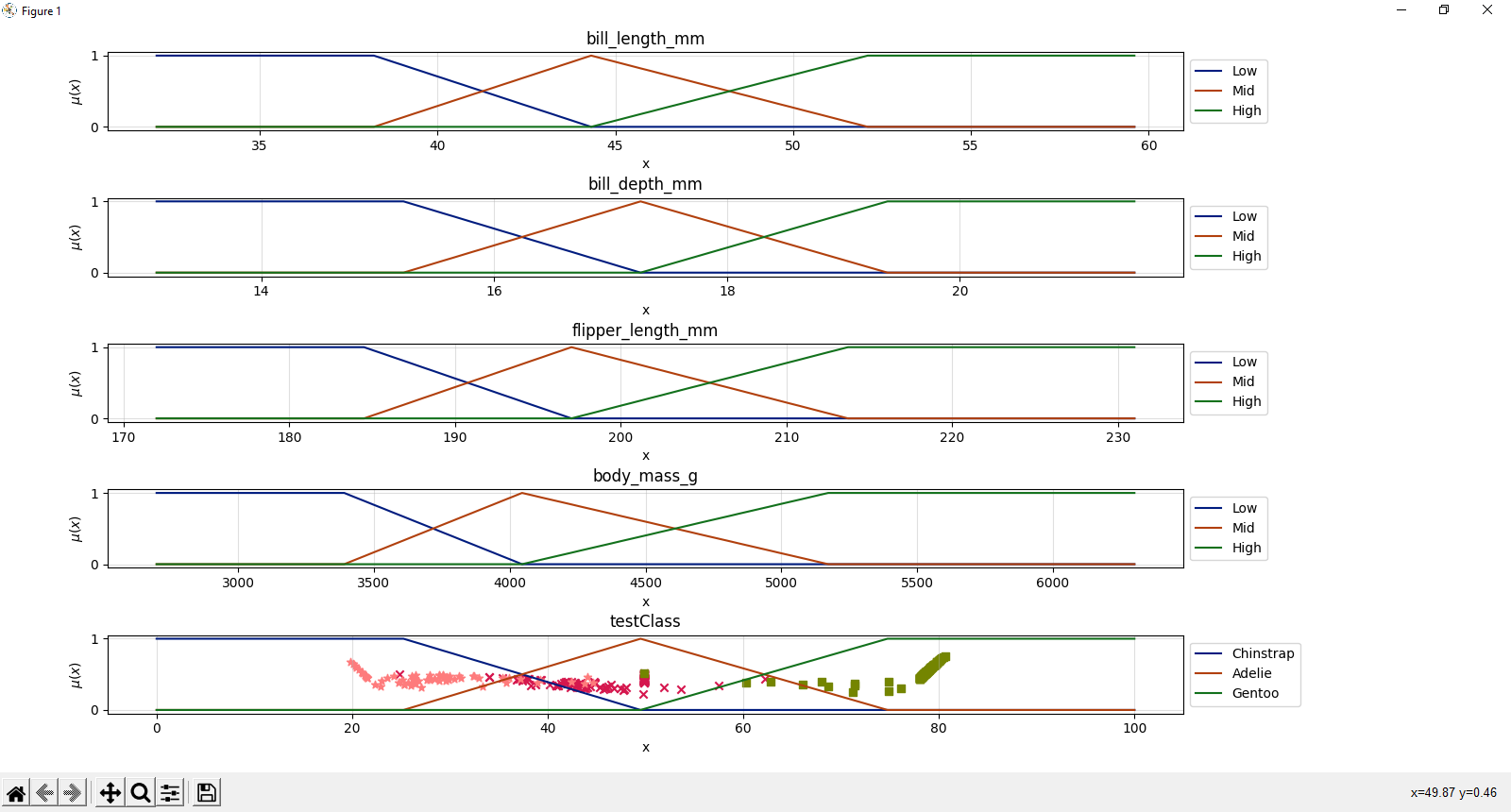


Рис. 4. После обучения: Графики 4-х функций принадлежности и результата

Рассмотрена зависимость результата классификации, от размера обучающей выборки до/после обучения. Для этого были сформированы обучающая и тестовая выборки в размере 200/144 (обучающая/тестовая). На рис. 5 представлена матрица противоречий до обучения системы, в табл. 1 приведены значения метрик качества.

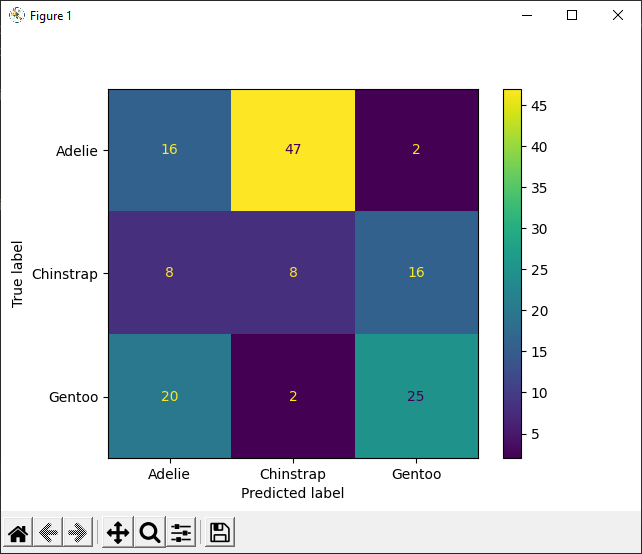


Рис. 5. До обучения: матрица противоречий – 200/144 (обучающая/тестовая)

Таблица 1

До обучения: 200/144 (обучающая / тестовая)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| Adelie | 0.36 | 0.25 | 0.29 | 65 |
| Chinstrap | 0.14 | 0.25 | 0.18 | 32 |
| Gentoo | 0.58 | 0.53 | 0.56 | 47 |
| avg accuracy | 0.34 | | | |

На рис. 6 представлена матрица противоречий после обучения системы, в табл. 2 приведены значения метрик качества.

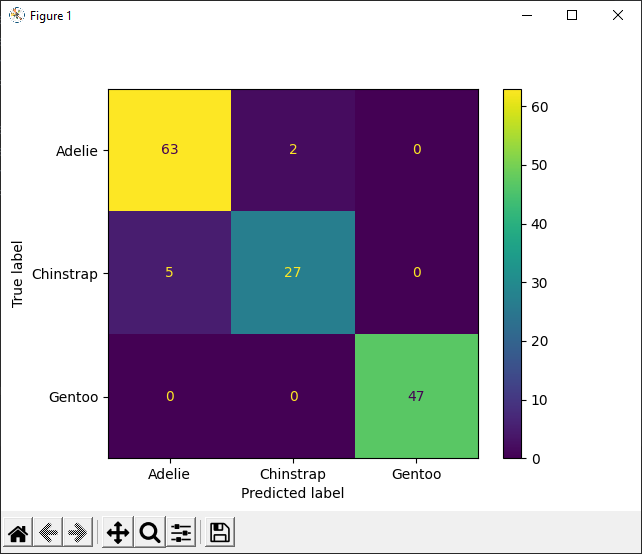


Рис. 6. После обучения: матрица противоречий – 200/144 (обучающая/тестовая)

Таблица 2

После обучения: 200/144 (обучающая / тестовая)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| Adelie | 0.93 | 0.97 | 0.95 | 65 |
| Chinstrap | 0.93 | 0.84 | 0.89 | 32 |
| Gentoo | 1.00 | 1.00 | 1.00 | 47 |
| avg accuracy | 0.90 | | | |

Для анализа влияния процесса обучения на результат классификации были изменены размеры обучающей/тестовой выборок. На рис. 7 представлена матрица противоречий до обучения системы на 250 примерах, в табл. 3 приведены значения метрик качества при размере тестовой выборки из 94 объектов.

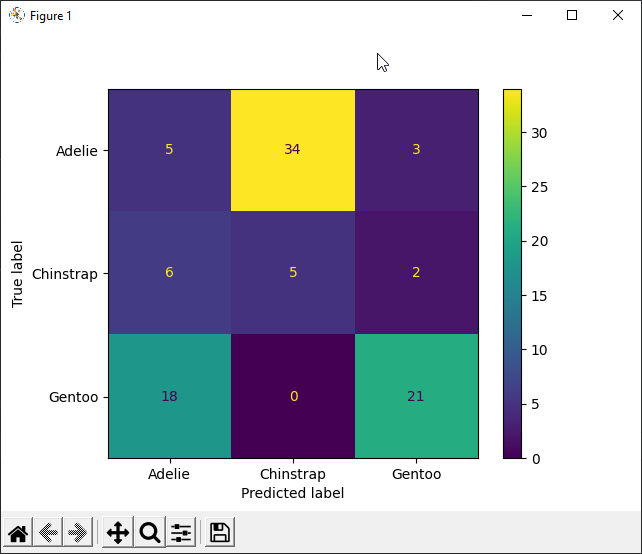


Рис. 7. До обучения: матрица противоречий – 250/94 (обучающая/тестовая)

Таблица 3

До обучения: 250/94 (обучающая / тестовая)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| Adelie | 0.17 | 0.12 | 0.14 | 42 |
| Chinstrap | 0.13 | 0.38 | 0.19 | 13 |
| Gentoo | 0.81 | 0.54 | 0.65 | 39 |
| avg accuracy | 0.33 | | | |

После обучения результат матрицы противоречий представлен на Рис. 8 и анализ метрик в табл. 4.

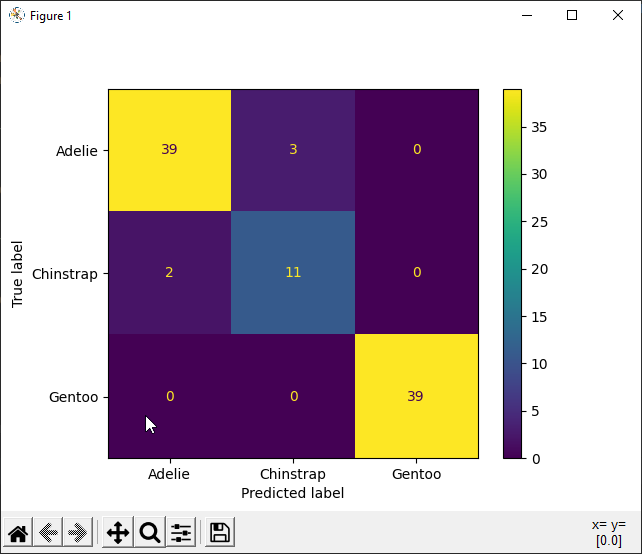


Рис. 8. После обучения: матрица противоречий – 250/94 (обучающая/тестовая)

Таблица 4

После обучения: 250/94 (обучающая / тестовая)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **f1-score** | **support** |
| Adelie | 0.95 | 0.93 | 0.94 | 42 |
| Chinstrap | 0.79 | 0.85 | 0.81 | 13 |
| Gentoo | 1.00 | 1.00 | 1.00 | 39 |
| avg accuracy | 0.93 | | | |

**6. Обсуждение результатов**

Анализ диаграммы предельного распределения (рис. 1) показал, что виды Adelie и Chinstrap достаточно связаны и их классификация может быть затруднительной. Особенно это заметно для признаков body\_mass\_g и bill\_depth\_mm.

На рис. 2 приведен результат проверки базы знаний на полноту при помощи предиката слабейшего постусловия. Результат «Valid» показывает истинность предиката *Q =>WP(S,R),* а, следовательно, полноту базы знаний.

Применяя полученную базу знаний до обучения на рис. 3 можно увидеть, что классы не определены и значения в результатах перемешаны. Отсутствует группирование объектов по классам. Но, после обучения на рис. 4 Можно увидеть, что данные сгруппировались и каждый класс находится в границах своей функции принадлежности. Следует отметить, что на графике присутствуют одновременно все 344 объекта.

Далее рассмотрим влияние размера обучающей выборки на результат классификации и качество экспертной системы. Для этого была сформирована обучающая и тестовая выборки в размере 200/144 (обучающая/тестовая). До обучения на рис. 5 представлена матрица противоречий и анализ метрик в табл. 1. По результатам можно заметить, что точность достаточно низкая 34 %. После обучения по рис. 6 и анализу метрик в табл. 2 можно заметить, что точность системы возросла до 90 %, а количество неверных ответов в матрице противоречий значительно уменьшилось.

Далее была сформирована обучающая и тестовая выборки в размере 250/94 (обучающая/тестовая). По рис. 6 и табл. 3 также можно заметить низкую точность 33 %. Но, после обучения точность возросла до 93 % как и остальные показатели в табл. 4.

Из анализа таблиц 1, 2, 3, 4 видно, что увеличение объема обучающей выборки, позволяет получить приемлемую точность классификации для выборок, а так же настроить вектор консеквентов для решения задачи классификации.

К ограничениям применения данного метода, можно отнести ситуацию, когда объекты, которые необходимо классифицировать обладают сильно связанными признаками. На диаграмме предельного распределения (рис. 1) виды Adelie и Chinstrap и их признаки особенно body\_mass\_g и bill\_depth\_mm сильно связаны и их классификация может быть затруднительной. Если сильно связанных объектов много в задаче, которую необходимо решить, то классификация может быть затруднительной.

Недостатком данного подхода является относительно большой объем базы правил. При увеличении числа признаков  для объекта *xi ∈ X*, количество лингвистическим переменным *k1, k2,…,kL* так же возрастает, что ведёт к росту числа правил. В дальнейшем необходимо применить методы редукции базы правил, для уменьшения числа правил. Что позволит применять данный подход к задачам, где объекты обладают большим числом признаков.

**7. Выводы**

1. В работе была построена продукционная модели представления знаний объединяющая модели Мамдани и Такаги-Сугено-Канг с функцией в качестве весового коэффициента для консеквента правила. Приведены методы для фазизфикации и дефазификации продукционной модели.

2. Разработан алгоритм автоматического построения базы правил. Матрица антецедентов создаётся при помощи признаков объектов из некоторого конечного множества и Декартового произведения. Предложено метод формирования вектора консеквентов на основе обучающей выборки. Антецеденты продукционной системы правил были представлены в виде матрицы, а консеквенты правил в виде вектора столбца.

3. Разработан подход для проверки корректности продукционных правил, доказательство полноты системы правил произведено при помощи логики Хоара, метода резолюций, предиката Q=>WP(S,R), формальной логики исчисления первого порядка и программного обеспечения Simplify использующего вариацию метода резолюций.

4. Выбраны следующие метрики для оценки качества полученной нечёткой базы правил: accuracy, precision, recall, f1-score.

5. Созданный подход применён для решения задачи классификации. Проведена проверка корректности и оценка качества полученной нечёткой базы правил. Выполнен анализ результатов вычислительного эксперимента.

**Конфлікт інтересів**

Автори декларують, що не мають конфлікту інтересів стосовно даного до-слідження, в тому числі фінансового, особистісного характеру, авторства чи іншого характеру, що міг би вплинути на дослідження та його результати, представлені в даній статті.

**Литература:**

1. Zadeh L.A., Abbasov A.M., Yager R.R., Shahbazova S.N., Reformat M.Z. (2014). Recent developments and new directions in soft computing. Eds. Studfuz 317, Springer, 3-10. doi: <https://doi.org/10.1007/978-3-319-06323-2>

2. Терновой М. Ю., Штогрина Е. С. (2015). Формальная спецификация свойств баз нечетких знаний Мамдани на основе метаграфа. Вісник ХНУ імені В. Н. Каразіна. Серія: Математичне моделювання. Інформаційні технології. Автоматизовані системи управління, Вип. 27, 157-171.

3. Кривуля Г.Ф., Шкиль А.С., Кучеренко Д.Е. (2013). Анализ корректности продукционных правил в системах нечеткого логического вывода с использованием квантовых моделей. АСУ и приборы автоматики: всеукр. межвед. науч.-техн. сб. – Х.: Изд-во ХНУРЭ, Вып. 165. 42-53

4. Darwiche A., Marquis P. (2002). A Knowledge Compilation Map. Journal of Artificial Intelligence 17. 229-264. doi: <https://doi.org/10.1613/jair.989>

5. Sugiuraj A., Riesenhuberj M., Kosekij Y. (1993). Comprehensibility Improvement of Tabular Knowledge Bases. AAAI-93 Proceedings, ISBN 978-0-262-51071-4, 716-721

6. Sugiuraj A., Kosekij Y. (1995). Comprehensibility Improvement of Tabular Knowledge Bases. Journal of the Japanese Society for Artificial Intelligence, Volume 10, Issue 4, 628-635. doi: <https://doi.org/10.11517/jjsai.10.4_628>

7. Кондратенко Ю.П., Козлов А.В. (2019). Генерація баз правил нечітких систем на основі модифікованих мурашиних алгоритмів. Міжнародний наук.-техн. журнал «Проблеми керування та інформатики», № 2, 59-79.

8. Желдак Т. А., Коряшкіна Л. С. (2020). Нечіткі множини в системах управління та прийняття рішень. навч. посіб. М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка» – Дніпро : НТУ «ДП», ISBN 978-960-350-726-2, 222- 227.

9. Hoare, C.A.R. (1969) An axiomatic basis for computer programming Communications of the ACM, 12 (10), 576–580 and 583.  
doi: <https://dl.acm.org/doi/10.1145/363235.363259>

10. Gries D. (1987). The Science of Programming (Monographs in Computer Science). Springer New York Inc, ISBN-13:978-0387964805, 108-114.

11. Dijkstra E. W. (1975) Guarded commands, nondeterminacy and formal derivation of programs. Commun. ACM, 18, ISSN 0001-0782, 453–457.  
doi: <https://doi.org/10.1145/360933.360975>.

12. Simplify The Extended Static Checker software. Available at: <https://www.kindsoftware.com/products/opensource/escjava2>

13. Хижа А.Л., Высокопоясный И.Г. (2017). Автоматическая проверка семантической правильности решений задач по программированию. Питання прикладної математики і математичного моделювання. Випуск 17, ISSN 2074-5893, 234-246. doi: 10.15421/321727

14. Gorman K.B., Williams T.D., Fraser W.R. (2014). Ecological sexual dimorphism and environmental variability within a community of Antarctic penguins (genus Pygoscelis). PLoS ONE 9(3):e90081, 1-14. doi: <https://doi.org/10.1371/journal.pone.0090081>

**References:**

1. Zadeh L.A., Abbasov A.M., Yager R.R., Shahbazova S.N., Reformat M.Z. (2014). Recent developments and new directions in soft computing. Eds. Studfuz 317, Springer, 3-10. doi: <https://doi.org/10.1007/978-3-319-06323-2>

2. Ternovoy M. YU., Shtohryna E. S. (2015). Formalʹnaya spetsyfykatsyya svoystv baz nechetkykh znanyy Mamdany na osnove metahrafa. Visnyk KHNU imeni V. N. Karazina. Seriya: Matematychne modelyuvannya. Informatsiyni tekhnolohiyi. Avtomatyzovani systemy upravlinnya, Vyp. 27, 157-171.

3. Krivulya G.F., Shkil' A.S., Kucherenko D.Ye. (2013). Analiz korrektnosti produktsionnykh pravil v sistemakh nechetkogo logicheskogo vyvoda s ispol'zovaniyem kvantovykh modeley. ASU i pribory avtomatiki: vseukr. mezhved. nauch.-tekhn. sb. – KH.: Izd-vo KHNURE, Vyp. 165. 42-53

4. Darwiche A., Marquis P. (2002). A Knowledge Compilation Map. Journal of Artificial Intelligence 17. 229-264. doi: <https://doi.org/10.1613/jair.989>

5. Sugiuraj A., Riesenhuberj M., Kosekij Y. (1993). Comprehensibility Improvement of Tabular Knowledge Bases. AAAI-93 Proceedings, ISBN 978-0-262-51071-4, 716-721

6. Sugiuraj A., Kosekij Y. (1995). Comprehensibility Improvement of Tabular Knowledge Bases. Journal of the Japanese Society for Artificial Intelligence, Volume 10, Issue 4, 628-635. doi: <https://doi.org/10.11517/jjsai.10.4_628>

7. Kondratenko YU.P., Kozlov A.V. (2019). Heneratsiya baz pravyl nechitkykh system na osnovi modyfikovanykh murashynykh alhorytmiv. Mizhnarodnyy nauk.-tekhn. zhurnal «Problemy keruvannya ta informatyky», № 2, 59-79.

8. Zheldak T. A., Koryashkina L. S. (2020). Nechitki mnozhyny v systemakh upravlinnya ta pryynyattya rishenʹ. navch. posib. M-vo osvity i nauky Ukrayiny, Nats. tekhn. un-t «Dniprovsʹka politekhnika» – Dnipro : NTU «DP», ISBN 978-960-350-726-2, 222- 227.

9. Hoare, C.A.R. (1969) An axiomatic basis for computer programming Communications of the ACM, 12 (10), 576–580 and 583.  
doi: <https://dl.acm.org/doi/10.1145/363235.363259>

10. Gries D. (1987). The Science of Programming (Monographs in Computer Science). Springer New York Inc, ISBN-13:978-0387964805, 108-114.

11. Dijkstra E. W. (1975) Guarded commands, nondeterminacy and formal derivation of programs. Commun. ACM, 18, ISSN 0001-0782, 453–457.  
doi: <https://doi.org/10.1145/360933.360975>.

12. Simplify The Extended Static Checker software. Available at: <https://www.kindsoftware.com/products/opensource/escjava2>

13. Khizha A.L., Vysokopoyasnyy I.G. (2017). Avtomaticheskaya proverka semanticheskoy pravil'nosti resheniy zadach po programmirovaniyu. Pitannya prikladnoí̈ matematiki í matematichnogo modelyuvannya. Vipusk 17, ISSN 2074-5893, 234-246. doi: 10.15421/321727

14. Gorman K.B., Williams T.D., Fraser W.R. (2014). Ecological sexual dimorphism and environmental variability within a community of Antarctic penguins (genus Pygoscelis). PLoS ONE 9(3):e90081, 1-14. doi: <https://doi.org/10.1371/journal.pone.0090081>

|  |  |
| --- | --- |
| **Англ.** | **Укр.** |
| Natalia Guk  Doctor Sciences in Physics and Mathematics, Professor, Head of the Department  Department of Computer Technology  Oles Honchar Dnipro National University  D. Yavornitsky ave., 35, Dnipro, Ukraine, 49000  Е-mail: [natalyguk29@gmail.com](mailto:natalyguk29@gmail.com)  Contact tel: (067) 565-60-71  ORCID: [http://orcid.org/0000-0001-7937-1039](http://orcid.org/0000-0001-7937-1039?lang=ru)  Number of articles in national databases – 60  Number of articles in Number of articles in international databases – 7 (Scopus)  ResearchGate: <https://www.researchgate.net/profile/Natalia-Guk> | Гук Наталія Анатоліївна  Доктор фізико-математичних наук, професор, завідувачка кафедри  Кафедра комп’ютерних технологій  Дніпровський національний університет імені Олеся Гончара  пр. Д. Яворницького, 35, м. Дніпро, Україна, 49000  Е-mail: [natalyguk29@gmail.com](mailto:natalyguk29@gmail.com)  Контактний тел.: (067) 565-60-71  ORCID: [http://orcid.org/0000-0001-7937-1039](http://orcid.org/0000-0001-7937-1039?lang=ru)  Кількість статей в загальнодержавних базах даних – 60  Кількість статей в міжнародних базах даних – 7 (Scopus)  ResearchGate: <https://www.researchgate.net/profile/Natalia-Guk> |
| Danylo Yehoshkin  Postgraduate student  Department of Computer Technology Oles Honchar Dnipro National University  D. Yavornitsky ave., 35, Dnipro, Ukraine, 49000  Е-mail:  [KnightDanila@i.ua](mailto:KnightDanila@i.ua)  Contact tel.: (050) 481-75-56  ORCID: <https://orcid.org/0000-0002-0937-4733>  Number of articles in national databases – 2  Number of articles in international databases – 0 | Єгошкін Данило Ігорович  Аспірант  Кафедра комп’ютерних технологій  Дніпровський національний університет імені Олеся Гончара  пр. Д. Яворницького, 35, м. Дніпро, Україна, 49000  Е-mail: [KnightDanila@i.ua](mailto:KnightDanila@i.ua)  Контактний тел.: (050) 481-75-56  ORCID: <https://orcid.org/0000-0002-0937-4733>  Кількість статей в загальнодержавних базах даних – 2  Кількість статей в міжнародних базах даних – 0 |

Єгошкін Данило Ігорович

Відділення Нової пошти №1 (вулиця Наукова, 1, Дослідне, Дніпропетровська область, 52071)

Тел: *(050) 481-75-56*

Якщо Ви вважаєте, що даний рукопис підходить для опублікування у «Східно-Європейському журналі передових технологій», то ми пропонуємо наступних рецензентів:

**Кісельова Олена Михайлівна**

чл.-кор. НАН України,

професор, доктор фізико-математичних наук,

декан факультету прикладної математики

Дніпровського національного університету

імені Олеся Гончара

пр. Д.Яворницького, 35, корпус 3,

м. Дніпро, Україна, 49000

Е-mail: leljka6@gmail.com

Контактний тел.: (050)-580-47-98

ORCID: http://orcid.org/0000-0003-4303-1707

**Чумаченко Дмитро Ігорович**

канд. техн. наук, доцент на кафедрі 304

Національного аерокосмічного університету

ім. М. Є. Жуковського «Харківський авіаційний інститут»

вул. Чкалова, 17, к. 135

м. Харків, Україна, 61070

Е-mail: dichumachenko@gmail.com

Контактний тел.: +38 (057) 788-40-06

ORCID: <https://orcid.org/0000-0003-2623-3294>