

# РЕЦЕНЗИЯ

**върху дисертационен труд за придобиване  
на образователната и научна степен „доктор”  
по научна специалност „Компютърни системи и технологии”  
Област на висшето образование 5. Технически науки  
Професионално направление 5.3. Комуникационна и  
компютърна техника**

Автор на дисертационния труд:

**Станислав Константинов Попов**

Тема на дисертационния труд:

**Обобщени мрежи и Data mining**

Рецензент:

**доц. д-р инж. Олимпия Николаева Роева**

- 1. Актуалност на разработвания в дисертационния труд проблем в научно и научно-приложно отношение. Степен и нива на актуалност на проблема и конкретните задачи, разработени в дисертацията**

Извличането на знания от данни (Data mining) е процес, при който се използват различни техники и алгоритми от области като машинното обучение, разпознаването на образи, статистиката, изкуствения интелект и др. Информацията, която трябва да бъде обработена за вземането на правилни и обосновани решения, става все повече. В същото време изискванията към оперативността и качеството на вземаните решения стават все по-високи. Всичко това изисква оптимално използване на наличните данни за вземане на решения и получаване на допълнително знание. Налице е необходимостта от усъвършенстване както на алгоритмите за обработка на данни, така и на програмното обезпечаване.

В контекста на гореспоменатото, разглежданият проблем в дисертационния труд – да се изследват различни процеси от теорията на извличането на знания от данни (Data Mining) чрез моделирането им с помощта на обобщени мрежи и програмна реализация – е актуален както в научно, така и в научно-приложно отношение.

Докторантът си поставя следните задачи, които дават възможност за качествено решаване на разглеждания проблем:

1. Да се анализират методите за извличане на закономерности от данни чрез алгоритми за Data Mining процеси.
2. Да се анализират алгоритми за клъстеризация на данни.
3. Разработване на обобщеномрежов модел на MapReduce.
4. Разработване на обобщеномрежов модел на стохастичен Expectation-Maximization алгоритъм.
5. Разработване на обобщеномрежов модел на Deep Learning невронна мрежа.
6. Разработване на обобщеномрежови модели за клъстерен анализ: по CLIQUE (клъстеризация в QUES), йерархична клъстеризация, клъстеризация по метода STING.
7. Програмна реализация на MapReduce и K-means алгоритмите.

## **2. Степен на познаване състоянието на проблема и творческа интерпретация на литературния материал**

Дисертационният труд е с обем от 113 страници и списък с 139 литературни източници. Около 34% от литературните източници са от последните 10 години, а станалите са основополагащи и/или значими в разглежданата област.

В литературния обзор е представена същността на процеса на извличане на знания от данни, както и на основните градивни блокове при този процес. Направен е кратък преглед на теорията на обобщените мрежи, както и на развитието ѝ във времето.



### **3. Съответствие на избраната методика на изследване с поставената цел и задачи на дисертационния труд**

Избраната методика и възприетите подходи съответстват на нивото на съвременната теория и постижения, и съответстват на дефинираните в дисертационния труд задачи. Разработени са нови обобщеномрежови модели на алгоритми за *Data Mining* и клъстеризация, които, от една страна, допълват съществуващите модели, а от друга, засилват взаимовръзката между теорията на обобщените мрежи и извличането на знания от данни.

### **4. Кратка аналитична характеристика на естеството и оценка на достоверността на материала, върху който се градят приносите на дисертационния труд.**

Дисертационният труд е добре структуриран и логически последователен съгласно дефинираните задачи за решаване. Съдържа 3 глави – една обзорна глава и две глави с резултати на докторанта.

Глава 1 е обзорна и представя състоянието на разглеждания проблем по литературни данни.

В Глава 2 от дисертационния труд са представени три обобщеномрежови модела на процеси, свързани с извличане на знания от данни – MapReduce, Stochastic Expectation-Maximization и Deep Learning Neural Network. Алгоритъмът MapReduce е реализиран в среда MATLAB. Представени са графични резултати. Анализирани са предимствата на представените модели, както и възможностите им за приложение. Например, обобщеномрежовият модел на процеса на MapReduce може да бъде използван за изследване и наблюдение, а чрез използването на апарата на обобщените мрежи, описанието на Stochastic Expectation-Maximization алгоритъм допринася за по-задълбочено вникване в начина на действие на алгоритъма. Моделът на Deep Learning Neural Network е подходящ за отправна точка при моделирането на други типове невронни мрежи.

В Глава 3 са представени обобщеномрежови модели на алгоритми за клъстеризация – клъстерен анализ, използващ CLIQUE; йерархичен клъстерен анализ и клъстерен анализ, използващ STING. Направена е графична симулация на k-means алгоритъма в MATLAB.

Разработените обобщеномрежови модели способстват за анализ и наблюдение на процесите на клъстеризация и дават възможност за последващи подобрения на реалните процеси за клъстеризация.

#### **5. Научни и научно-приложни приноси на дисертационния труд.**

Докторантът формулира следните научни и научно-приложни приноси:

Научни приноси:

- Създадени са обобщеномрежови модели на действието на следните алгоритми за извличане на знания: MapReduce, Deep Learning невронна мрежа и Stochastic Expectation-Maximization.
- Разработени са обобщеномрежови модели на процеса на клъстерен анализ, както следва: по CLIQUE (клъстеризация в QUEst), йерархична клъстеризация, клъстеризация по метода STING. Впоследствие изпълнението на алгоритмите може да бъде категоризирано чрез обща обобщена мрежа.

Научно-приложни приноси:

- На базата на изчислителните възможности на MapReduce алгоритъма и създадения обобщеномрежов модел са реализирани две тествания. Те са свързани с моделиране на вероятност чрез логистична регресия в MATLAB. Използваните данни са за пациенти на болница.
- В MATLAB е реализирана графична програмна демонстрация на действието на k-means алгоритъма, като по този начин може да се направи паралел между него и йерархичната клъстеризация.

Според мен научните приноси следва да се разглеждат като научно-приложни, а научно-приложните като приложни.

Приемам два научно-приложни приноси, както следва:

1. Разработени са три обобщеномрежови модела на алгоритми за извличане на знания – MapReduce алгоритъм, Stochastic Expectation-Maximization алгоритъм и Deep Learning Neural Network.
2. Разработени са три обобщеномрежови модела на алгоритми за клъстеризация – клъстерен анализ, използващ CLIQUE; йерархичен клъстерен анализ и клъстерен анализ, използващ STING.



Два приложни приноса:

3. MapReduce алгоритъма е реализиран и тестван в среда MATLAB. Тестванията са свързани с моделиране на вероятност чрез логистична регресия като са използваните реални данни за пациенти.
4. В среда MATLAB е реализирана графична програмна демонстрация на действието на  $k$ -means алгоритъма. Реализацията дава възможност да се направи паралел между  $k$ -means алгоритъма и йерархичния клъстерен анализ.

#### **6. Преценка на публикациите по дисертационния труд.**

Докторантът представя 6 публикации по дисертационния труд: 3 публикации в специализирани издания реферирани в Scopus с импакт ранг, две от които на Springer; 1 публикация индексирани в IEEE Xplore и 2 публикации в българско списание.

В посочените публикации са представени основните приноси на С. Попов.

Представени са и данни за 6 цитирания на публикациите по дисертационния труд.

#### **7. Оценка на съответствието на автореферата с изискванията за изготвянето му, както и на адекватността на отразяване на основните положения и приносите на дисертационния труд.**

Авторефератът правилно отразява съдържанието на дисертацията и дава представа за разглежданите проблеми, както и за приносите на дисертационния труд.

#### **8. Мнения, препоръки и бележки**

Според мен представеният програмен код на реализираните в MATLAB алгоритми би трябвало да се изнесе в края на дисертационния труд като Приложение. В основния текст да останат резултатите (графични и числени), за които липсват коментари и дискусия.

Липсва и глава Заключение на дисертационния труд.

Забелязват се някои технически грешки и неточности, например в библиографията, но тъй като пропуски от такова естество по никакъв начин не омаловажават качествата на дисертационния труд, не считам за необходимо да ги разглеждам.

В дисертационния труд (на стр. 98) се споменава, че *„Изграденият обобщеномрежов модел може да способства за постигането на по-нататъшни подобрения на реалния процес.“* Какви могат да бъдат тези подобрения или в какви насоки ще се търсят? Как или с какво описанието с обобщеномрежови модел води/може да доведе до идеи за усъвършенстване на алгоритмите?

Въпреки посочените забележки и коментарът ми относно приносите на дисертационния труд, Станислав Попов представя резултати с научно-приложен характер, които представляват принос в областта и са публикувани в реномирани специализирани научни издания и са представени на международни научни форуми.

#### **9. Заключение с ясна положителна или отрицателна оценка на дисертационния труд**

Моето мнение е, че докторантът демонстрира добри познания в конкретната област и научен потенциал за реализация на оригинални идеи.

На основата на гореизложено, считам че дисертационният труд на Станислав Попов отговаря на изискванията на Закона за развитие на академичния състав в Република България (ЗРАСРБ), Правилника за прилагане на ЗРАСРБ, Правилника за условията и реда за придобиване на научни степени и за заемане на академични длъжности в Университет „Проф. д-р Асен Златаров“. Постигнатите резултати ми дават основание да предложа на уважаемото научно жури да присъди образователната и научна степен „доктор“ на Станислав Попов.

01.12.2020 г.

София

Подпис заличен  
Чл.2 от ЗЗЛД

(доц. д-р О. Роева)