

# СТАНОВИЩЕ

от доц. д-р Николай Тодоров Кочев

на дисертационен труд за присъждане на образователната и научна степен 'доктор'

в област на висше образование *4. Природни науки, математика и информатика*

професионално направление *4.2. Химически науки*

докторска програма *Аналитична химия*

**Автор:** Стефка Руменова Начкова

**Тема:** Компютърни методи за интерпретация на  $^{13}\text{C}$  ЯМР спектри на природни съединения

**Научен ръководител:** Проф. д.х.н. Пламен Николов Пенчев

## **Общо представяне на процедурата и докторанта**

Представеният от Стефка Руменова Начкова комплект материали е в съответствие с Чл.36 (1) от Правилника за развитие на академичния състав на ПУ и включва всичките необходими документи.

## **Актуалност на тематиката**

Дисертационният труд разглежда актуална тематика в областта на ефективното извличане на информация от молекулни спектри. Отразени са хеометрични подходи и съвременни информационни подходи за data mining.

## **Познаване на проблема**

Докторант Стефка Начкова показва задълбочено познаване на състоянието на проблематиката на дисертационния труд. В литературния обзор подробно са разгледани и дискутирани подходите за интерпретация на ЯМР спектри чрез библиотечно търсене, експертни системи и хеометрични подходи. Докторантката е анализирала много изчерпателно компютърно базираните системи за интерпретация на спектри, като са разгледани основните им характеристики и принципите на работа. Докторант Стефка Начкова е направила адекватна оценка на състоянието на проблемите за интерпретация на ЯМР спектри, въз основа на които са дефинирани основната цел и практическите задачи на дисертацията.

## **Методика на изследването**

Методиката на изследванията е основана на комбинирането на множество и разнообразни похвати: компютърна обработка на спектри и структури, създаване на нови спектрални

библиотеки, работа с различни специализирани софтуерни програми по химична информатика за обработка на химични линейни нотации и структурни представяния, софтуери за библиотечно търсене, статистически анализ на структурна и спектрална информация, линейна регресия и хеометрични методи за класификация и моделиране. При работата с всеки софтуерен инструмент, докторант Начкова е демонстрирала познаване на детайлите в дълбочина и умело използване за целите на дисертацията. Осъществен е безкомпромисен и ефективен статистически анализ, валидиране и последваща обоснова на получените резултати. Впечатляващо е използването на методи за симулация на спектри с цел автоматично изчистване на грешки (data curation) в спектралните колекции, като по този начин са открити и отстранени редица неточности в данните от различни литературни източници.

### **Характеристика и оценка на дисертационния труд и приносите**

Докторант Начкова е извършила мащабна и разнообразна научно-изследователска работа. Осъществена е изчерпателна литературна справка, в която много оптимизирано и точно са разгледани основните методи за компютърно подпомогнато разкриване на структури на химични съединения, като е формулирана актуална цел на дисертационния труд и подходящи задачи за постигането на целта. Тематиката на дисертацията изисква задълбочени мултидисциплинарни познания, които докторант Стефка Начкова несъмнено е демонстрирала.

Спазени са добрите практики за тестване и статистическо валидиране на създадените модели.

Оформлението на дисертацията като цяло се характеризира с оптимален обем, ясна логика и последователност на изложените текстове, добре формулирани мисли и тези и подходящо приготвени фигури и таблици. Тези характеристики правят дисертацията приятна за четене и възприемане на съдържанието.

Дисертационният труд има следните **научни и научно-приложни приноси**:

- Съставена е нова библиотека РНУСНЕМ от 1000  $^{13}\text{C}$  ЯМР спектъра на фитосъединения;
- Съставени са няколко вероятностни функции и са предложени процедури за оценка на надеждността на подструктурите, извлечени при интерпретационно търсене на природни съединения в библиотека РНУСНЕМ;
- Предложена е мярка за оценка на информационното съдържание на подструктурите, извлечени в резултат на интерпретационно библиотечно търсене;
- Намерени са оптималните толеранси на съвпадение на сигналите за търсене по подобие и идентичност на фитосъединения в различните библиотеки;

- Предложена е мярка за оценка на ефективността на търсене по подобие, която отчита подобие на получените структури в хитсписька, средното и максимално структурно подобие на библиотеката.

### **Забележки по съдържанието и оформлението на дисертацията:**

Наблюдават се съвсем незначителен брой технически и пунктоационни грешки.

- На страница 53 е използвана променливата  $i$  за обозначаване на праг, който се променя от 0.000 до 1.000 със стъпка 0.001. Смятам, че използването на такава променлива е неподходящо (и малко подвеждащо), защото обикновено в математиката и информатиката,  $i$  се използва за обозначаване на целочислени променливи, индексирани на суми и др.

- в таблица 8 (на стр. 74) описанието на съдържанието на колоните е обръквощо, тъй като се премесват понятията “праг на надеждност“, измерван в проценти и “прецизност“ също измервана в проценти. Текстът от предходния абзац разяснява донякъде съдържанието на таблицата.

- На фигура 14 липсват обозначения за случаите: а, б, в, г, което затруднява разчитането на фигурата.

### **Въпроси към докторанта:**

- В дисертацията (на стр. 47) споменавате, че работите със стойност на  $m.p.c. = 6$ . Защо сте избрали тази стойност? Правили ли сте изследване за влиянието на тази стойност върху резултатите от алгоритмите?

- На фигури 15 и 16 се наблюдава изместване на оптималната стойност на  $f$  от 0.6 към 0.85 при намаляване на праговата надеждност от 99 към 90%. Имате ли обяснение на какво се дължи това изместване на оптималната стойност на  $f$ ?

- На фигура 31, медианата на разпределението на подобие на всички двойки структури за фингърпринти V500 е силно изместена спрямо медианата на разпределението за *Klekota-Roth count*. Аналогично съотношение се наблюдава и при сравнение на фигури 28 и 29. Имате ли обяснение за този резултат? Според тези резултати, фингърпринтите на Вармуца отчитат наличие на вътрешно структурни корелации в много по-голяма степен. Може ли да се направи извод, кой от двата фингърпринта е по-подходящ?

### **Преценка на публикациите и личния принос на докторанта**

Научните съобщения, свързани с дисертацията, включват две публикации, едната от които е в специализирано международно списание с импакт-фактор 0.47. Досега не са забелязани цитати на публикациите. Личният принос на докторантката в изготвянето на резултатите и публикуването на всяко от научните съобщения е съществен.

## ЗАКЛЮЧЕНИЕ

Дисертационният труд *съдържа научни и научно-приложни резултати, които представляват оригинален принос в науката* и отговарят на изискванията на Закона за развитие на академичния състав в Република България (ЗРАСРБ), Правилника за прилагане на ЗРАСРБ и съответния Правилник на ПУ „Паисий Хилендарски“. Представените материали и дисертационни резултати **съответстват** на специфичните изисквания на Химическия факултет за присъждане на образователната и научна степен 'доктор', приети във връзка с Правилника на ПУ за приложение на ЗРАСРБ.

Дисертационният труд е с актуална тематика, като са извършени огромни по обем научни изследвания. Дисертационният труд показва, че докторант Стефка Начкова **притежава** задълбочени теоретични знания и професионални умения по научна специалност Аналитична химия като **демонстрира** качества и потенциал за самостоятелно провеждане на научно изследване.

Поради гореизложеното, убедено давам своята **положителна оценка** за проведеното изследване, представено в дисертационния труд, автореферата, постигнатите резултати и приноси, и **предлагам на почитаемото научно жури да присъди образователната и научна степен ‘доктор’** на Стефка Руменова Начкова в област на висше образование: *4. Природни науки, математика и информатика*, професионално направление *4.2. Химически науки* докторска програма *Аналитична химия*.

25.04.2018 г.

Изготвил становището: .....

доц. д-р Николай Кочев