

Облікова картка ДіР



I. Загальні відомості

Державний обліковий номер: 0225U000370

Державний реєстраційний номер: 0123U103184

Особливі позначки: відкрита

Дата реєстрації: 10-01-2025

II. Етап виконання ДіР

Номер етапу: 2

Назва етапу: Розробити методи негладкої оптимізації для задач кластеризації та задач статистичної обробки результатів спостережень

Початок етапу: 02.2024

Закінчення етапу: 12.2024

Вид звітного документа: Остаточний звіт

III. Відомості про виконавця ДіР

Повне найменування юридичної особи (або ПІБ фізичної особи): Інститут кібернетики імені В. М. Глушкова Національної академії наук України

Код за ЄДРПОУ: 05417176

Місцезнаходження: проспект Академіка Глушкова, буд. 40, м. Київ, 03187, Україна

Форма власності:

Сфера управління: Національна академія наук України

Ідентифікатор ROR: Не застосовується

Розмір організації:

Телефон: 380445262008, 380445264178

IV. Відомості про співвиконавців ДіР

V. Відомості про замовника ДіР

Повне найменування юридичної особи (або ПІБ фізичної особи): Національна академія наук України

Код за ЄДРПОУ: 00019270

Місцезнаходження: вул. Володимирська, буд. 54, м. Київ, 01601, Україна

Форма власності:

Сфера управління:

Ідентифікатор ROR: Не застосовується

Розмір організації:

Телефон: 380442343243

VI. Джерела, напрями та обсяги фінансування ДіР

Підстава для проведення ДіР: 34 - договір (замовлення) з центральним органом виконавчої влади, академією наук (головними розпорядниками бюджетних коштів на проведення НДДКР)

Напрямок фінансування: 2.1 - фундаментальні дослідження

Джерела фінансування

7713 - кошти держбюджету

Код програмної класифікації видатків і кредитування (КПКВК): 6541030

Фактичний обсяг фінансування (тис. грн.): 130.000
--

VII. Відомості про ДіР

Назва роботи українською:

Розробити алгоритми негладкої оптимізації для кластеризації та статистичної обробки даних

Назва роботи англійською:

To develop non-smooth optimization algorithms for clustering and statistical data processing

Реферат українською:

В роботі досліджено задачу розміщення складів, обґрунтовано умови сумісності системи обмежень задачі та розглянуто її варіанти залежно від умов балансу, які визначають виродженість та невиврожденість системи обмежень. Для задачі знаходження параметрів лінійної регресійної моделі з L1-регуляризацією та критерієм суми модулів нев'язок у степені $p \in [1, 2]$ побудовано алгоритм `emlmpg`. Результати тестових експериментів демонструють час роботи програми, робастність розв'язків задачі при $p \approx 1$ та якість відновлення параметрів моделі за наявності лінійної залежності між факторами. Сформульовано модель булевого лінійного програмування для задачі класифікації та описано метод обчислення нижніх оцінок цільової функції моделі на основі негладкої двоїстої задачі, отриманої за допомогою лагранжевої релаксації частини обмежень початкової задачі. Побудовано суперградієнтний алгоритм феєрівського типу для її розв'язання. Наведено результати тестових експериментів з використанням двоїстого підходу та проведено порівняння з результатами солвера Gurobi, та алгоритму `kmeans++`. Досліджено задачу знаходження нижньої оцінки коефіцієнту достовірності Гутмана для задачі навчального тестування та пов'язану з нею задачу нелінійного програмування, сформульовану як задачу максимізації лінійної функції при негладких обмеженнях на мінімальне власне число симетричної матриці, яка зводиться до безумовної максимізації негладкої функції. Побудовано тестовий приклад для її розв'язання за допомогою r -алгоритму, проведено порівняння одержаних результатів з результатами Флетчера. Для моделі двоетапної транспортної задачі з обмеженнями на невідомі потреби споживачів обґрунтовано необхідні та достатні умови сумісності системи обмежень. Сформульовано частковий випадок вихідної задачі, коли нижні та верхні границі на потреби споживачів співпадають, та її модифікацію, в якій деякі споживачі вилучаються з плану перевезень, а продукція транспортується лише до фіксованої кількості споживачів.

Реферат англійською:

Investigated in the paper is storage location problem, substantiated consistency conditions of constraint system of the problem, considered its variants depending on the balance conditions that determine degeneracy and non-degeneracy of constraint system. For finding parameters of linear regression model with L1-regularization and criterion of sum of residuals moduli powered to $p \in [1, 2]$, the `emlmpg` algorithm is constructed. Test experiments results demonstrate program running time, robustness of solutions of the problem at $p \approx 1$, and quality of model parameters restoration if there is a linear dependence between the factors. A Boolean linear programming model is formulated for classification problem and method for calculating lower estimates of model's objective function based on non-smooth dual problem obtained via Lagrangian relaxation of some constraints of the initial problem is described. A Feyer-type supergradient algorithm is constructed for its solution. Experiments results using the dual approach are presented and compared with Gurobi and `kmeans++` results. Problem of finding lower bound of the Gutman reliability coefficient for educational testing problem and related nonlinear programming problem, formulated as linear function maximizing problem under non-smooth constraints on minimum eigenvalue of symmetric matrix, which is reduced to unconditional maximization of a non-smooth function, are investigated. Test case is constructed for its solution using r -algorithm, and results obtained are compared with Fletcher's ones. For the model of a two-stage transportation problem with constraints on unknown consumer demands, necessary and sufficient conditions for constraint system consistency are substantiated. A partial case of the original problem is formulated, when the lower and upper bounds on consumer needs coincide, and its modification, in

which some consumers are removed from transportation plan, and products are transported only to fixed number of consumers.

Індекс УДК: 519.8 , 519.7:004.8

Коди тематичних рубрик: 27.47.19, 27.47.23

Керівники роботи

Власне Прізвище Ім'я По-батькові: Стовба Віктор Олександрович

Науковий ступінь: д.філософ

Вчене звання:

Ідентифікатор ORCID ID:

Додаткова інформація:

VIII. Наукова (науково-технічна) продукція (НТП)

Назва НТП українською: Моделі задач кластеризації та негладкі моделі задач статистичної обробки результатів спостережень

Назва НТП англійською: Clustering problem models and non-smooth models of statistical processing of observation results

НТП, яку передбачалося створити:

Причини, через які НТП не було створено:

Отримані результати: Методи, теорії

Галузь застосування: Приватні та державні підприємства сфери логістики, штучного інтелекту, медицини та біології, енергетики, сільського господарства тощо.

Реєстраційний номер картки технології:

Опис НТП: Розроблені моделі та методи можна використовувати для розв'язання задач кластеризації та класифікації, обробки результатів спостережень, відновлення параметрів регресійних моделей з наявністю аномалій та залежностей між ознаками, знаходження оцінок коефіцієнту достовірності в задачах навчального тестування, розв'язання двоетапної транспортної задачі з обмеженнями на невідомі потреби споживачів.

Соціально-економічна спрямованість НТП: Поліпшення якості життя та здоров'я населення, ефективності діагностики та лікування хворих

Вплив НТП на довкілля:

Впровадження НТП: Не впроваджено

Практична реалізація НТП

Початок етапу: 01.2025

Закінчення етапу: 12.2025

Споживачі продукції: Приватні та державні підприємства сфери логістики, штучного інтелекту, медицини та біології, енергетики, сільського господарства тощо.

Перспективні ринки: Україна

Характер співробітництва з інвестором

Потрібний обсяг інвестицій, тис. грн.:

Права, що надаються інвестору після завершення роботи:

Наявність бізнес-плану:

Техніко-економічне обґрунтування:

Потенціальний обсяг продажу, тис. грн.:

Очікуваний термін окупності (років):

Додаткова інформація:

IX. Бібліографічний опис

Stetsyuk P., Stovba V., Senko I., Chaikovsky I. Using of Ellipsoid Method and Linear Regression with L1-Regularization for Medical Data Investigation Proceedings of the 4th International Workshop of IT-professionals on Artificial Intelligence (ProfIT AI 2024). Cambridge, MA, USA, September 25-27, 2024. P. 210-221. <https://ceur-ws.org/Vol-3777/paper14.pdf>

Stetsyuk, P., Stovba, V., Korablov, M. Using of Ellipsoid Method for Finding Linear Regression Parameters with L1-Regularization Communications in Computer and Information Science. Vol. 2225. Springer, Cham. P. 350-362. https://doi.org/10.1007/978-3-031-73417-5_27

Stovba V., Khomiak O. Two-Stage Transportation Problem with Unknown Consumer Demands: Consistency Aspects Proceedings of 3rd International Conference on Problems of Logistics, Management and Operation in the East-West Transport Corridor (PLMO 2024), Baku, Azerbaijan, May 15-17, 2024.

Х. Заклучні відомості

Керівник юридичної особи

Сергієнко Іван Васильович
д.ф.-м.н.

Перелік осіб-виконавців

Корабльов Микола Миколайович
(122)

**Відповідальний за підготовку
облікових документів**

Стовба Віктор Олександрович

Телефон

+38 (050) 583-89-27

Реєстратор

**Керівник відділу УкрІНТЕІ, що є
відповідальним за реєстрацію наукової
діяльності**

