

## ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ ЗА ДОПОМОГОЮ РЕГРЕСІЙНОГО АНАЛІЗУ

*Сомряков Б. О., студент факультету  
комп'ютерних наук спеціальності  
122 «Комп'ютерні науки»*

*Чорноморський національний університет імені Петра Могили, Україна*

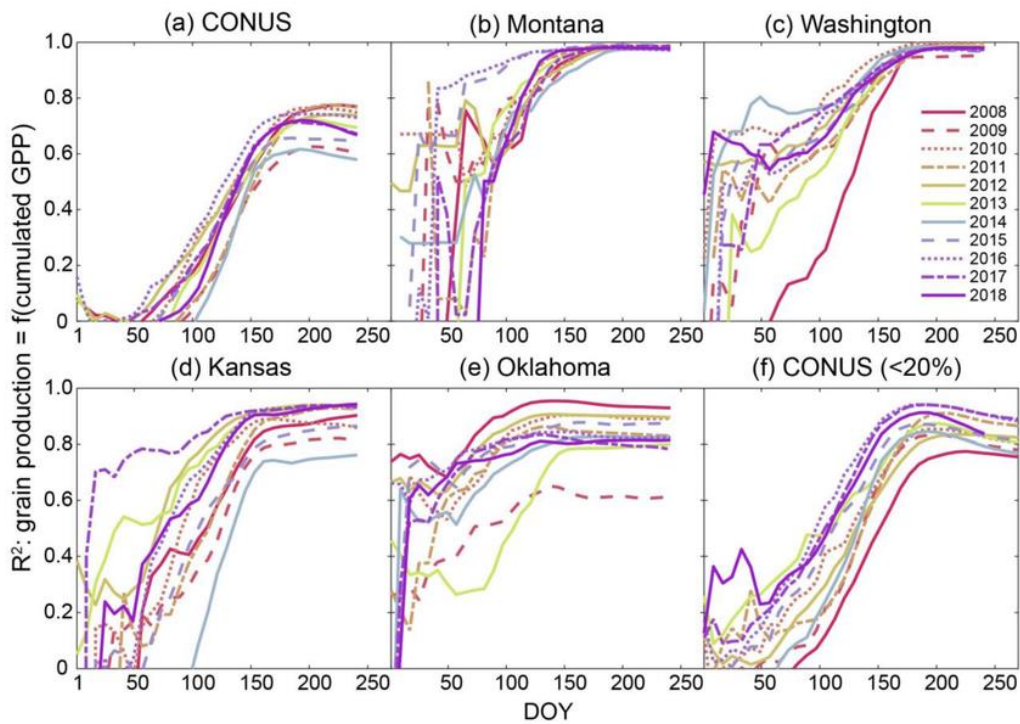
Регресійний аналіз — це статистичний метод, який може перевірити гіпотезу про те, що змінна залежить від однієї або декількох інших змінних. Далі регресійний аналіз може дати оцінку величини впливу зміни однієї змінної на іншу. Ця остання особливість, звичайно, важлива для прогнозування майбутніх значень [1].

Регресивний аналіз є ключовим для прогнозування врожайності, оскільки він дозволяє виявити взаємозв'язки між врожаєм і факторами, такими як клімат, стан ґрунту чи супутникові показники (наприклад, GPP), обробляючи великі обсяги даних для побудови точних моделей. Завдяки цьому можна передбачити врожайність, оптимізувати використання ресурсів, покращити планування аграрних робіт і забезпечити ефективність прийняття рішень.

На рисунку 1 зображено графіки, що відображають залежність між показником врожайності та накопиченим валовим первинним виробництвом (GPP) для різних регіонів США в різні роки.

На рисунку 1 представлено, як залежність між накопиченим валовим первинним виробництвом (GPP) і врожайністю змінюється в залежності від року і регіону. У всіх регіонах, наприклад, у Montana та Kansas, відзначається чітке збільшення коефіцієнта детермінації ( $R^2$ ) з часом, що свідчить про поліпшення точності моделі прогнозування врожайності впродовж року. Згідно з графіками, дані з різних років (2008-2018) показують різну ефективність моделі в кожному році: у деяких випадках, як, наприклад, у 2008 році, прогнозування є менш точним, тоді як у інші роки модель має високе значення  $R^2$ , що вказує на сильнішу кореляцію між GPP і врожайністю.

Інтерпретуючи графіки для різних регіонів, можна побачити, що деякі регіони, як Washington і Kansas, мають стабільніші і точніші моделі прогнозування врожайності, оскільки лінії  $R^2$  для цих регіонів показують постійне зростання чи високе значення протягом року. Водночас для інших регіонів, таких як Oklahoma, деякі роки мають нижче значення  $R^2$ , що вказує на більшу варіативність і меншу точність прогнозів. Це може свідчити про різні агротехнічні умови чи кліматичні фактори, що впливають на ефективність моделей у цих областях.



**Рисунок 1 – Залежність між показником врожайності та накопиченим валовим первинним виробництвом (GPP) для різних регіонів США [2]**

*Джерело: Spatiotemporal Changes of Winter Wheat Planted and Harvested Areas, Photosynthesis and Grain Production in the Contiguous United States from 2008–2018. URL: [https://www.researchgate.net/figure/The-prediction-skill-of-the-linear-regression-models-that-predict-county-level-crop-grain\\_fig10\\_351367432](https://www.researchgate.net/figure/The-prediction-skill-of-the-linear-regression-models-that-predict-county-level-crop-grain_fig10_351367432)*

Інтерпретуючи графіки для різних регіонів, можна побачити, що деякі регіони, як Washington і Kansas, мають стабільніші і точніші моделі прогнозування врожайності, оскільки лінії  $R^2$  для цих регіонів показують постійне зростання чи високе значення протягом року. Водночас для інших регіонів, таких як Oklahoma, деякі роки мають нижче значення  $R^2$ , що вказує на більшу варіативність і меншу точність прогнозів. Це може свідчити про різні агротехнічні умови чи кліматичні фактори, що впливають на ефективність моделей у цих областях.

Рисунок 2 ілюструє алгоритм використання лінійної регресії для прогнозування врожайності сільськогосподарських культур.

Спочатку використовуються дані про врожай (Crop Yield Dataset), які формують навчальний набір даних. Цей набір використовується для побудови моделі лінійної регресії, яка на основі вхідних ознак (Input Features) формує гіпотезу.

Далі, на основі цієї гіпотези, модель генерує прогнозований результат (Predicted Output), який потім порівнюється з фактичним результатом (Actual Output). Під час навчання модель коригує свою лінію регресії так, щоб помилка передбачення була якомога меншою, тобто мінімізується різниця між фактичними значеннями і прогнозованими значеннями. На

рисунок також показана лінія регресії, яка найкраще підходить для даних, де помилка прогнозування є мінімальною.

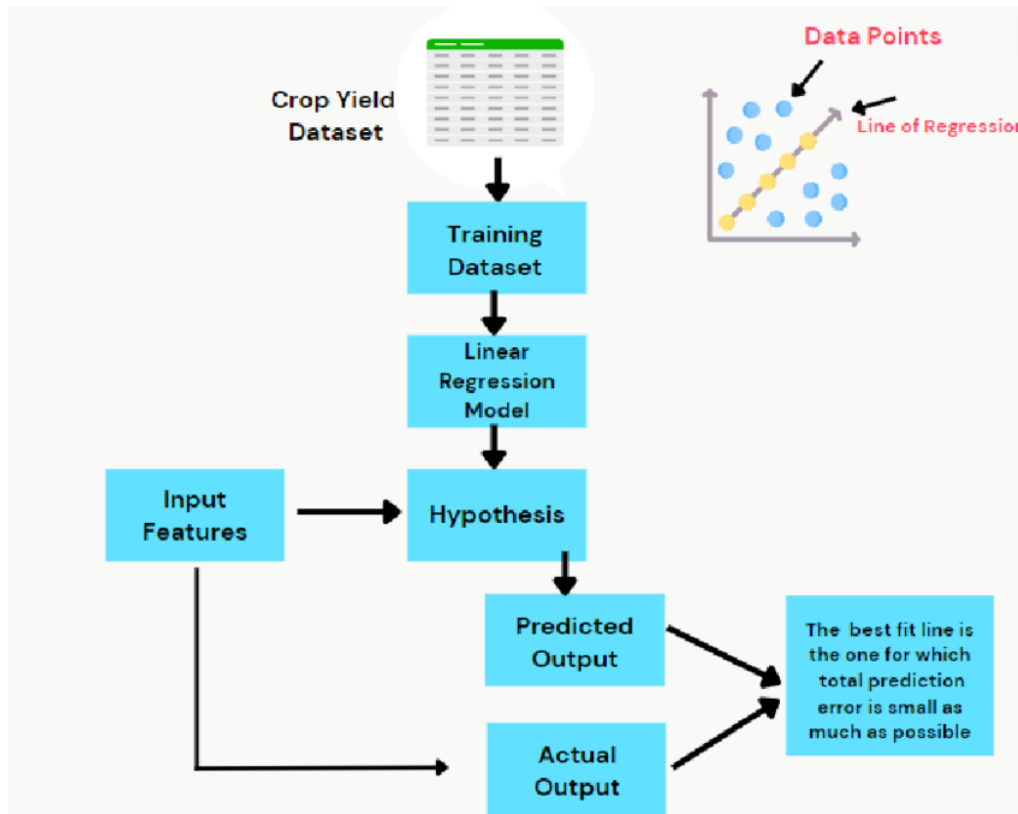


Рисунок 2 – Використання лінійної регресії для прогнозування врожайності сільськогосподарських культур [3]

Джерело: *Optimizing Crop Yield Forecasts Using Quantum Machine Learning Techniques with High-Dimensional Soil and Weather Data*. URL: [https://www.researchgate.net/figure/Linear-Regression-model-predicting-crop-Yield-flow-diagram\\_fig2\\_385723039](https://www.researchgate.net/figure/Linear-Regression-model-predicting-crop-Yield-flow-diagram_fig2_385723039)

Отже, регресійний аналіз є важливим інструментом для аграріїв, оскільки дозволяє прогнозувати врожайність на основі історичних даних і сучасних умов, оптимізувати використання ресурсів і підвищувати ефективність виробництва. Використовуючи дані, такі як валове первинне виробництво (GPP), моделі лінійної регресії створюють точні прогнози, мінімізуючи різницю між фактичними та прогнозованими значеннями врожаю.

## Список використаних джерел

1. Рівняння регресії. URL:

[https://ukrayinska.libretexts.org/Статистика/Прикладна\\_статистика/Книга%3A\\_Статистика\\_бізнесу\\_\(OpenStax\)/13%3A\\_Лінійна\\_регресія\\_та\\_кореляція/13.04%3A\\_Рівняння\\_регресії](https://ukrayinska.libretexts.org/Статистика/Прикладна_статистика/Книга%3A_Статистика_бізнесу_(OpenStax)/13%3A_Лінійна_регресія_та_кореляція/13.04%3A_Рівняння_регресії)

2. Spatiotemporal Changes of Winter Wheat Planted and Harvested Areas, Photosynthesis and Grain Production in the Contiguous United States from 2008–2018. URL:

[https://www.researchgate.net/figure/The-prediction-skill-of-the-linear-regression-models-that-predict-county-level-crop-grain\\_fig10\\_351367432](https://www.researchgate.net/figure/The-prediction-skill-of-the-linear-regression-models-that-predict-county-level-crop-grain_fig10_351367432)

3. Optimizing Crop Yield Forecasts Using Quantum Machine Learning Techniques with High-Dimensional Soil and Weather Data. URL: [https://www.researchgate.net/figure/Linear-Regression-model-predicting-crop-Yield-flow-diagram\\_fig2\\_385723039](https://www.researchgate.net/figure/Linear-Regression-model-predicting-crop-Yield-flow-diagram_fig2_385723039)