

УДК 004.94:004.032.26

DOI: 10.31319/2519-8106.1(38)2018.129020

Л.І. Коротка, к.т.н., доцент (доцент), korliv@hotmail.com

ДВНЗ «Український державний хіміко-технологічний університет», м. Дніпро

АНАЛІЗ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ В ЗАДАЧАХ ОПТИМІЗАЦІЇ ТЕХНОЛОГІЇ ЕНЕРГОКОНДЕНСОВАНИХ СИСТЕМ

У роботі розглядаються декілька архітектур нейронних мереж в оптимізації технології енергоконденсованих систем. Досліджено можливості використання нейромережових модулів у задачах апроксимації. Описано головні етапи побудови нейромережового компоненту. Розроблено і розглянуто різні моделі архітектур та проведено з ними чисельні експерименти. Зроблено аналіз отриманих результатів та обрано раціональну архітектуру нейромережового компоненту для подальшого використання у предметній області.

Ключові слова: нейронні мережі, технологія енергоконденсованих систем.

In this work the different architectures of neural networks in technology optimization of energy condensed systems. Modules studied neural network approximation tasks. Describes the basic steps to create neuro network component. Created and reviewed various models of architecture, numerical experiments with them. Conducted an analysis of the results and selected rational architecture neuro network component for its further use in the subject area.

Keywords: neural networks, technology of energy-condensed systems.

Постановка проблеми

Визначення оптимальних параметрів в хімічній технології нерозривно пов'язано з моделюванням промислових процесів. Відмова від промислового застосування тротилу та тротилових вибухових речовин є світовою тенденцією розвитку гірничодобувних технологій. Вирішення цієї проблеми можливе тільки за рахунок створення і впровадження нових безпечних технологій нітратних енергоконденсованих систем (ЕКС), які складаються з невибухових компонентів і характеризуються високим рівнем безпечності поводження та застосування.

В технології безтритилових вибухових речовин однією з проблем є визначення технологічних параметрів. Складність визначення оптимальних параметрів моделювання технологічного процесу полягає в тому, що основний вхідний параметр — міцність гранул сировини (гранульованої аміачної селітри (АС)) залежить від ряду факторів, визначення яких ускладнене або неможливе: технологічні параметри грануляції АС, природа та концентрація використаних модифікуючих добавок, наявність антизлежувачів, термін та температурні умови зберігання АС після виготовлення тощо.

Аналіз останніх досліджень та публікацій

В рамках вирішення цієї проблеми в Україні була розроблена технологія вискоефективних гранульованих ЕКС марки Україніт-АНФО [1, 2], яка передбачає модифікацію гранульованої аграрної аміачної селітри з метою підвищення пористості гранул при збереженні їх міцності. Гранули АС у процесі модифікації обробляють поризуючим водним розчином (0,2—0,8 % від маси АС) з подальшим видаленням внесеної вологи вакуумною сушкою в обертальному апараті періодичної дії. Збільшення пористості аміачної селітри при цьому реалізується за рахунок розкриття твердої оболонки гранули та термічних перетворень основного матеріалу гранули, а зміцнення — за рахунок створення «жорсткого» каркасу з важкорозчинних продуктів взаємодії речовини поризуючого розчину та промислових добавок АС, що вже містяться в гранулі. Промислове впровадження розробленої технології виявило необхідність корегування технологічних параметрів (кількості та концентрації поризуючого розчину) в залежності від механіко-міцностних характеристик гранул вхідної сировини (аграрної АС) [3].

Формулювання мети дослідження

Задача оптимізації технологічних параметрів передбачає моделювання процесу модифікації, що дозволить прогнозувати параметр відносної динамічної міцності гранул кінцевого продукту в залежності від характеристик вхідної сировини та технологічних параметрів. За-

уважимо, що основними вхідними параметрами є динамічна міцність гранул сировини, кількість поризуючого розчину та масова концентрація поризуючого розчину. За результатами досліджень було встановлено, що цільовий параметр міцності модифікованих гранул (вихідний параметр — Y) змінюється з часом та набуває оптимальних значень на 7—8 добу після виготовлення [4]. Границі, в яких він змінюється встановлено у результаті натурних експериментів: $Y = 97 + 99,8 \%$.

Складність вирішення поставленої мети полягає в тому, що вхідний параметр, в свою чергу, залежить від ряду факторів, як вже зазначалося, визначення яких в умовах технології ЕКС ускладнене або неможливе.

В залежності від параметрів технологічного процесу, накопичення заводською лабораторією протягом року (середня продуктивність по ЕКС 200—250 т/місяць) статистичних даних про відносну динамічну міцність гранул кінцевого продукту дозволили реалізувати моделювання процесу модифікації.

Отримання математичної моделі з використанням нейронних мереж є завданням описаного технологічного процесу. Метою дослідження — отримання раціональної архітектури нейронної мережі при оптимізації технології енергоконденсованих систем.

Виклад основного матеріалу

Застосування класичних методів отримання аналітичних формул, зокрема, множинного регресійного аналізу, повного факторного експерименту та ін., не виключає запропонованого у роботі підходу. У якості альтернативного підходу пропонується використання технологій обчислювального інтелекту, а саме: штучних нейронних мереж (НМ).

Постановка задачі оптимізації технології енергоконденсованих систем передбачає отримання залежності $Y = F(X)$, де $X = (x_1, x_2, x_3, x_4, x_5)$:

x_1 — динамічна міцність гранул вихідної аміачної селітри (АС), %;

x_2 — кількість поризуючого розчину від маси АС, % мас;

x_3 — концентрація поризуючого розчину, %;

x_4 — кількість речовини поризуючого розчину, який вноситься під час поризації, % мас;

x_5 — кількість часу після виготовлення гранул, доба.

При цьому, змінні мають наступні обмеження та умови, що визначаються технологією модифікації: $85 \leq x_1 \leq 100$; $0,2 \leq x_2 \leq 0,8$; $12,25 \leq x_3 \leq 50$; $y \geq 0$.

Математична постановка. На основі даних натурних експериментів є вибірка $\{(X^k; Y^k), k = \overline{1, N}\}$, вхідний вектор $X = \{x_1, x_2, x_3, x_4, x_5\}$ — описаний вище; вихідний вектор Y — міцність готової продукції (ЕКС), % мас; N — об'єм навчальної вибірки (НВ). Необхідно побудувати нейромережевий компонент для апроксимації функції багатьох змінних $Y = F(X)$.

В залежності від розв'язуваної задачі обирається відповідна топологія нейронної мережі. Сформульована задача є задачею апроксимації. Як відомо, з такими задачами успішно справляються слоїсті багат шарові нейронні мережі, таким чином, топологія НМ є визначеною. Базовий математичний апарат побудови нейронних мереж достатньо детально наведено, наприклад, у [7, 9].

Згідно до теореми про повноту та наслідків з неї маємо: нейронні мережі є універсальними апроксимуючими системами. Відповідно до базового математичного апарату теорії НМ буде розглянуто одно та двошарові мережі з логістичними функціями активації. У якості алгоритму навчання нейронної мережі буде використано відомий алгоритм зворотного поширення помилки.

Розробка моделей, методів, алгоритмів та програмних комплексів для неформалізованих задач на основі нейромережевих компонентів для оцінки та прогнозування деяких показників є достатньо актуальною. Технологія проектування нейромережевого компоненту представлено на рис. 1 [8].

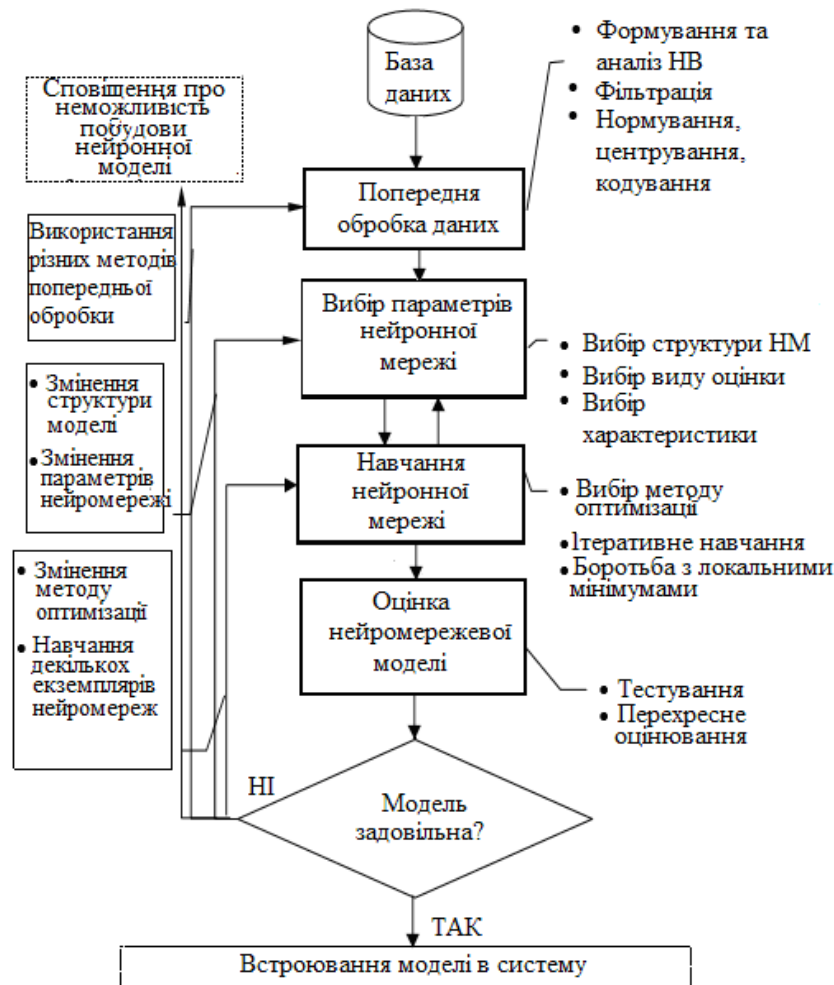


Рис. 1. Технологія проектування нейромережевого компоненту

Слід зауважити, що визначення раціональної архітектури конкретної нейромережевої моделі необхідно пройти кожен етап [6]: попередня обробка даних; вибір параметрів нейронної мережі; навчання НМ; оцінка якості моделі та її вбудовування у систему. Розглянемо кожен етап окремо.

Попередня обробка даних передбачає практично формування навчальної вибірки, а саме: необхідно залишити ті зразки, що адекватно та найбільш повно відображають зовнішнє середовище. Фактично цей етап було виконано на етапі попереднього експерименту при фіксуванні параметрів технологічного процесу. Слід зауважити, що в якості навчальної вибірки розглядаються результати пасивного натурного експерименту. Об'єм навчальних даних може доповнюватися, але тоді необхідне до навчання нейронної мережі, що з точки зору обчислювальних витрат виконати не складно.

Нормування вже сформованої навчальної вибірки є необхідною умовою та не вивіває додаткових пояснень. Для того, щоб перейти до наступного етапу слід зазначити, що об'єм навчальної вибірки N прямо пропорційно залежить від кількості вхідних даних та елементів прихованого шару/шарів мережі та обернено пропорційно залежить від заданої похибки НМ ε_{NM} [5, 6]. Як відомо, від презентативності навчальної вибірки залежить якість навчання нейронної мережі.

Наступний етап побудови нейромережевого компоненту передбачає вибір параметрів мережі. Зауважимо, що цей етап у багатьох випадках носить евристичний характер, тобто вибір параметрів залежить цілком від проектувальника НМ та рекомендацій щодо їх вибору [9]. З

одного боку це дозволяє експериментувати з великою кількістю архітектур/структур, а з іншого боку навіть ускладнює вибір.

Як зазначалося вище, у роботі розглянуто одно та двошарові нейронні мережі. Варіювалась кількість нейронів у прихованому шарі для одношарової мережі та кількість нейронів для двошарової НМ. Формули визначення кількості елементів у шарі є достатньо відомими [9].

У роботах [5, 9] відзначається, що проблема багатошарового персеプトону з одним прихованим шаром складається у тому, що нейрони прихованого шару взаємодіють на глобальному рівні. З одного боку, у складних задачах така взаємодія ускладнює підвищення якості апроксимації. З іншого боку, при наявності двох прихованих шарів процес апроксимації є більш керованим. Зауважимо, що останні твердження носять більш теоретичний, аніж практичний характер. Головним показником якості навчання мережі та вибору раціональної архітектури з наявних є контрольна похибка ϵ .

До параметрів нейронної мережі відносять: коефіцієнт навчання $\eta \in (0; 1)$, вибір функції активації (ФА) та початкові значення елементів матриці вагових коефіцієнтів. Слід зазначити, що вибір значення η покладається цілком на проектувальника нейромережевого компоненту. Як відомо, алгоритм зворотного поширення помилки використовує ідею методу градієнтного спуску (тобто відбувається рух по поверхні помилок в сторону мінімуму), тому якщо функція поверхні похибки має дуже «овражну» поверхню, то ймовірність попадання у западину при $\eta \rightarrow 0$ є достатньо високою, але гіршим є те, що завдяки такому коефіцієнту вибратися з неї не завжди вдається. З іншого боку велике значення $\eta \rightarrow 1$ може привести до ситуації, коли локальний екстремум буде пропущено. У зв'язку з цим у роботі коефіцієнт навчання варіювався, а саме для визначеності було розглянуто наступні його значення: $\eta = 0,1$; $\eta = 0,5$; $\eta = 0,9$. Отримані чисельні результати порівнювались згідно контрольної похибки мережі.

Матриці початкових значень вагових коефіцієнтів обирались згідно рекомендацій [9], тобто $w_{ij} \in [-0,3; 0,3]$.

Ще одним питанням, яке постає перед проектувальником нейронної мережі є вибір функції активації. Згідно наслідків з теореми про повноту вони повинні бути двічі диференційованими. Таким вимогам задовольняє велика кількість функцій, але дуже добре себе зарекомендували сигмоїдальні ФА (1)—(2), зокрема, логістична (рис. 2) та гіперболічний тангенс (рис. 3) [5].

$$f(\text{net}) = \frac{1}{1 + \exp(-c \cdot \text{net})}, \quad c = 1. \quad (1)$$

$$f(x) = a \cdot \text{th}(b \cdot x), \quad a = 1,7159; \quad b = 2/3. \quad (2)$$

Процес навчання нейронної мережі є достатньо відповідальним. Загальні підходи, які було застосовано мають відомий класичний підхід, який описано, наприклад, в [5, 6, 8, 9]. А саме, для навчання нейронної мережі береться 2/3 навчальної вибірки, а для наступного етапу тестування резервується остання її 1/3. Щодо вибору методу оптимізації, то, як зазначалося раніше, було обрано класичний алгоритм зворотного поширення помилки [7]. Автору відомо, що можливі модифікації цього алгоритму в залежності від розв'язуваної задачі, але такої необхідності застосування модифікованого алгоритму не виникало. Зауважимо, що для корегування синаптичних ваг використано дельта-правило (або правило Видроу-Хоффа), але необхідності застосування інерційного коефіцієнту не мало потреби. Як зазначається у [8], останні підходи можуть використовуватись для боротьби з локальними мінімумами.

Для покращення роботи ітеративного алгоритму вживають два режими подачі вхідних навчальних зразків: послідовний та пакетний. Вказані підходи є достатньо відомими, зазначимо, лише, що у роботі використано послідовний режим.

Останнім етапом побудови нейромережевого компоненту є перевірка його здатності узагальнювати, тобто етап тестування НМ. У цьому випадку використовується зарезервована третина зразків і саме по ній визначається контрольна похибка. Зауважимо, що вона може перевищувати задану похибку, але в незначному відсотковому відношенні.

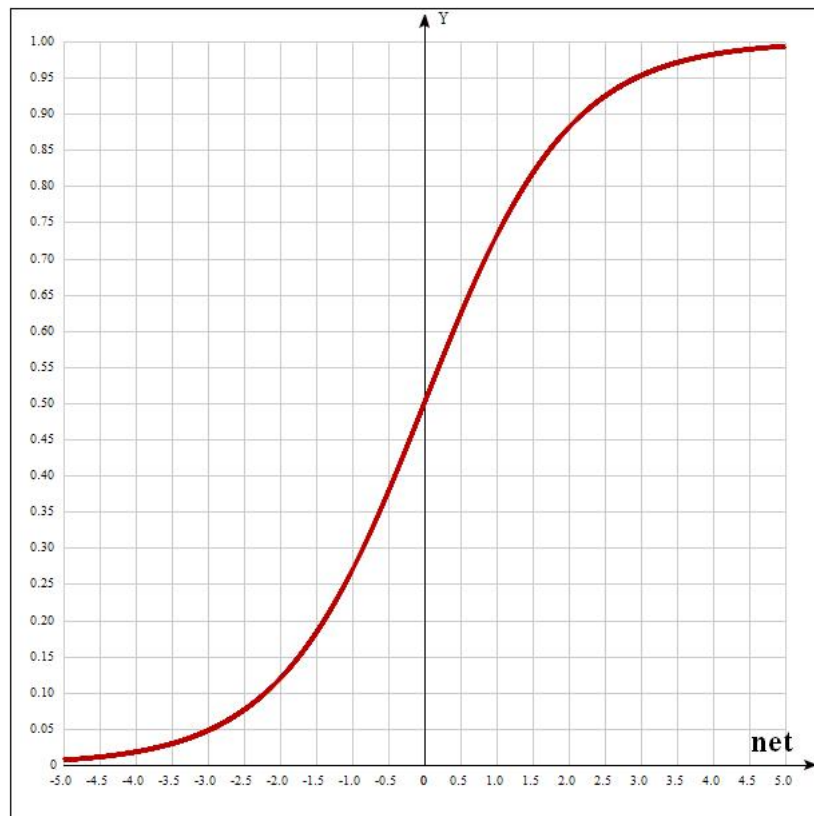


Рис.2. Логістична функція активації

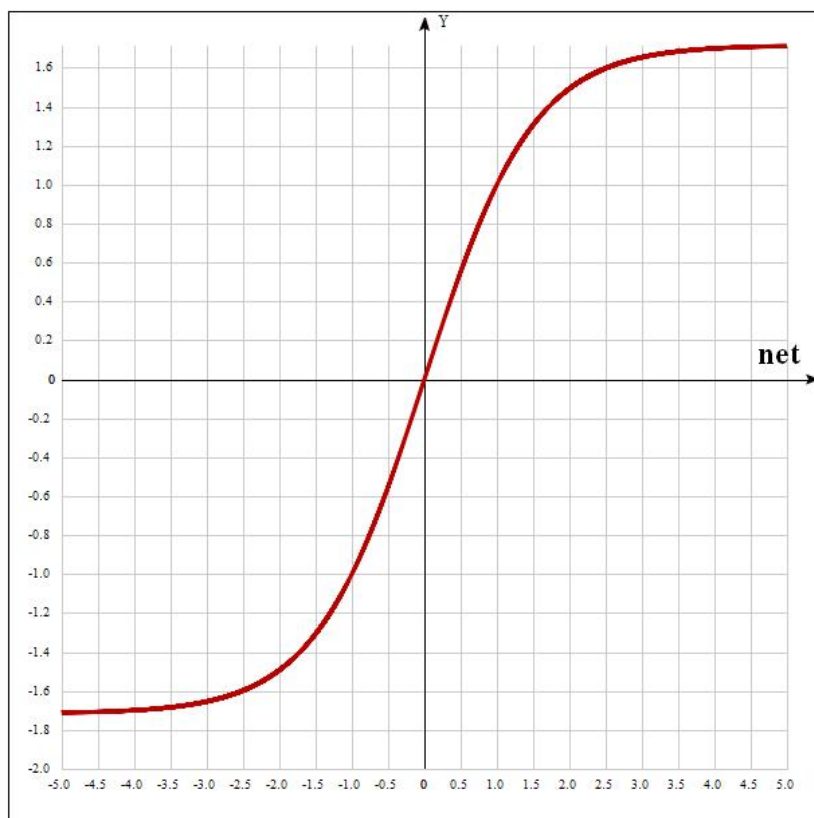


Рис.3. Функція активації — гіперболічний тангенс

Перед тим, як перейти до демонстрації чисельних експериментів, слід зазначити, що для кожної заданої похибки об'єм навчальної вибірки повинен бути необхідним. Автору відомо, що існує й верхня межа НВ. Для одношарової мережі з архітектурою [5-5-1] (позначимо її НМ(1)) з заданою похибкою $\varepsilon_{HM} = 0,1$ необхідний об'єм вибірки повинен складати $N = 100$ зразків, а відповідно для $\varepsilon_{HM} = 0,01$ складає $N = 1000$; а для НМ з архітектурою [5-9-1] (позначимо її НМ(2)) з заданою похибкою $\varepsilon_{HM} = 0,1$ необхідний об'єм НВ повинен складати $N = 140$ зразків, а відповідно для $\varepsilon_{HM} = 0,01$ складає $N = 1400$. Для двошарової НМ [5-5-5-1] (позначимо її НМ(3)) з заданою похибкою $\varepsilon_{HM} = 0,1$ необхідний об'єм вибірки повинен складати $N = 150$ зразків, а відповідно для $\varepsilon_{HM} = 0,01$ складає $N = 1500$.

Як зазначалося вище, у роботі використовується обмежена кількість зразків (хоча база даних може поповнюватись), але ця кількість навчальних зразків не є достатньою (усього до 100 зразків НВ). Тому у роботі виникла необхідність нарощування об'єму навчальної вибірки, а саме виконувалось так зване стохастичне перетасовування зразків. Таким чином, відбувалось нарощування навчальної вибірки. Для цього було написано окремий програмний модуль. Необхідність використання стохастичного нарощування об'єму НВ пояснюється тим, що якщо цього не робити, то якість процесу навчання НМ може значно знизитися. Можливе виникнення замкнених циклів і не тільки. Якщо ж кількість зразків буде надлишковою, тоді можливі ситуації так званого «перенавчання» мережі [5, 9], або навіть її «параліч».

Тому для визначення об'єму НВ та кількості нейронів у прихованих шарах використано відомі формули [5, 6].

Для попередження описаних ситуацій «перетренування» мереж було застосовано кросс-перевірки, що сприятливо відобразилось на якості навчання НМ різних архітектур.

Чисельні експерименти були проведені з різними моделями нейронних мереж. Слід зауважити, що ретельна підготовка вхідних зразків у достатній мірі сприяла тому, що різні мережі навчались за достатньо невелику кількість епох незалежно від функції активації, заданого кроку навчання та заданої похибки. Тільки у незначних випадках сумарна похибка тестової вибірки незначно перевищувала задану. Без обмеження міркувань у таблиці 1 наведено деякі

Таблиця 1. Результати чисельних експериментів

Тип архітектури НМ	Тип функції активації	Задана похибка	Контрольна похибка
НМ(1)	(1)	0,10	0,0692
	(1)	0,01	0,0464
	(2)	0,10	0,1098
	(2)	0,01	0,0399
НМ(2)	(1)	0,10	0,9680
	(1)	0,05	0,0543
	(2)	0,05	0,0305
	(2)	0,01	0,0399
НМ(3)	(1)	0,05	0,0503
	(2)	0,01	0,0156

результати чисельних експериментів для середнього значення коефіцієнту навчання НМ $\eta = 0,5$, хоча на практиці було проведено навчання нейронних мереж для різних (і крайніх у тому числі) значень η .

Після проведення чисельних експериментів з вибору тої чи іншої архітектури нейронної мережі цей процес можна представити (рис. 4).

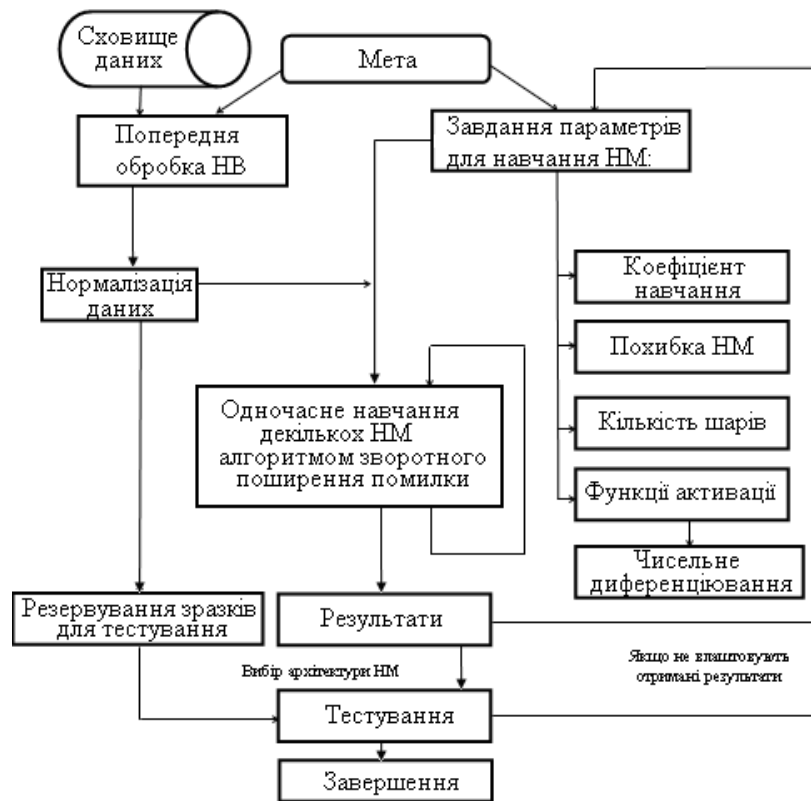


Рис. 4. Процес вибору раціональної архітектури НМ серед претендентів

Очевидно, що вибір параметрів нейронної мережі покладається повністю на проектувальника нейронної мережі. Після проведення чисельних експериментів за результатами навчання визначається краща з претендентів мережа.

Проектувальник визначає, що нейронна мережа має добрі здібності до узагальнення і саме її вбудовують у систему або використовують в подальшому.

Висновки

Аналіз отриманих результатів, дозволяє зробити наступні висновки. Розглянуті моделі мереж різних архітектур з логістичними функціями активації мають практично однакову узагальнюючу здібність. Саме тому, як зазначається [7], з такого різноманіття моделей-претендентів слід обирати мережу з простішою архітектурою. Тому було обрано навчену нейронну мережу НМ (1). Одержані результати нейромережевого моделювання впроваджені на виробництві гранульованих ЕКС в ПрАТ «Промвибух» (м. Запоріжжя). Застосування мережі дозволяє прогнозувати параметри ЕКС без приготування експериментальних порцій продукту [2].

Список використаної літератури

1. Коваленко, И.Л. Технология модифицирования аграрной аммиачной селитры в производстве энергоконденсированных систем / И.Л. Коваленко, Д.В. Киященко // Science and education a new dimension. Natural and technical sciences. 2015. III (8), Issue 73. P. 107-110
2. Пат. 106118 Україна, МПК С06В31/28, F42D3/04. Спосіб виготовлення гранульованих вибухових сумішей для пневматичного заряджання із аміачної селітри і рідкого палива / Купрін В.П., Купрін О.В., Іщенко М.І., Савченко М.В., Коваленко І.Л. – № у 2015 13112; заявл. 30.12.2015; опубл. 11.04.2016, Бюл. № 7.
3. Kovalenko, I. Cartridged and granulated explosive substances of grade Ukrainit for underground mines / I. Kovalenko, N. Stupnik, M. Korolenko, D. Kiyaschenko, A. Batareev // Metallurgical and

- mining industry. – 2016. – № 8. – Р. 59–64. Доступ до журналу: <http://www.metaljournal.com.ua/MMI-2016-No-8>.
4. Коваленко, І.Л. Вплив попередньої підготовки на вбираючу здатність і міцність гранул аміачної селітри марки Б // І.Л. Коваленко, В.П. Купрін // Наукові вісті НТУУ «КПІ». – 2014. – № 6 (98). – С. 110–114.
 5. Короткая, Л.И. Использование нейронных сетей при численном решении некоторых систем дифференциальных уравнений / Л.И. Короткая // Восточно-европейский журнал передовых технологий. – 2011. – № 3/4 (51). – С. 24–27.
 6. Коротка, Л.І. Функціональна підсистема раціонального вибору архітектури нейронної мережі / Л.І. Коротка // Вісник Херсонського національного технічного університету 3(62), Том I. (Фундаментальні науки) – 2017. С. 55–59.
 7. Круглов, В.В. Нечеткая логика и искусственные нейронные сети / В.В. Круглов, М.И. Дли, Р.Ю. Голунов. – М.: Физматлит, 2001. – 224 с.
 8. Пятковский, О.И. Интеллектуальные информационные системы. (Нейронные сети). Учебное пособие / Алт. гос. техн. Ун-т им. И.И. Ползунова. Барнаул: Изд-во АлтГТУ, 2010. – 125 с.
 9. Хайкин, С. Нейронные сети: полный курс, 2-е издание. пер. с англ. / Саймон Хайкин. – М.: Издательский дом "Вильямс", 2006. – 1104 с.