

## **Твір наукового характеру**

### **"Нейромережна модель процесу прогнозування характеристик програмного забезпечення на основі аналізу специфікацій"**

Процес розроблення програмного забезпечення (ПЗ) тісно пов'язаний з процесом аналізу та оцінювання значущих характеристик ПЗ: тривалості та моделі життєвого циклу, вартості, ефективності, складності, зручності використання, кросплатформеності, надійності та якості ПЗ. Критичний вплив на програмні проекти та на успішність їх завершення здійснюють питання, пов'язані із аналізом специфікації.

Існує велика кількість специфікацій та стандартів з їх написання. Основною, найбільш повною з точки зору визначення основних характеристик ПЗ, є специфікація вимог до ПЗ (SRS). Кількісні показники, які можна одержати із запропонованої стандартної структури SRS: термін виконання, кваліфікація виконавця(ів), кількість виконавців; розмір цільової аудиторії; кількість компонентів системи, розмір кожного компонента, повторюваність; вартість використовуваних операційних систем, баз даних, компіляторів; вартість специфічних засобів розробки; розмір і вартість документації для користувачів; залежність від інших програмних засобів (вартість, складність, надійність); перелік всіх функційних вимог, вартість, складність, надійність окремо взятої вимоги; складність та надійність окремо взятого алгоритму; вартість помилки, реакція на помилку; кількість, вартість інтерфейсів користувача; кількість, вартість, надійність, складність програмних інтерфейсів; кількість, вартість, надійність, складність апаратних інтерфейсів; кількість, вартість, надійність, складність інтерфейсів зв'язку і комунікацій; перелік всіх нефункційних вимог, вартість, складність, надійність окремо взятої вимоги. Якісні показники специфікації вимог до ПЗ: профілювання (цільова аудиторія); вміння користувачів працювати із аналогічними продуктами; складність операційних систем, баз даних, компіляторів; залежність між вимогами, суперечливість вимог; якість програмного забезпечення; рівень безпеки ПЗ.

*Визначення 1.* Кількісними показниками (КП) специфікації вимог до ПЗ або її розділів називатимемо корисні для визначення характеристик ПЗ показники специфікації, які мають точне кількісне значення (наприклад, "кількість виконавців - 7", "середня вартість помилки - 89\$" і т.і.).

*Визначення 2.* Якісними показниками специфікації вимог до ПЗ або її розділів називатимемо корисні для визначення характеристик ПЗ показники специфікації, які виражені лінгвістично (наприклад, "користувачі мають навички роботи із аналогічними продуктами" і т.і.).

*Визначення 3.* Експертними кількісними показниками (ЕКП) специфікації вимог до ПЗ або її розділів називатимемо корисні для визначення характеристик ПЗ показники, які у специфікації виражені лінгвістично, але після опрацювання групою експертів набули кількісного виразу (наприклад, вищенаведений якісний показник перетворений після опрацювання групою експертів у експертний кількісний показник "середня оцінка навичок роботи користувачів із аналогічними продуктами складає 3,8 (за 5-бальною шкалою)" і т.і.).

Залежність кожної основної характеристики ПЗ від кількісних та експертних кількісних показників специфікації вимог до ПЗ відображено на рис.1.

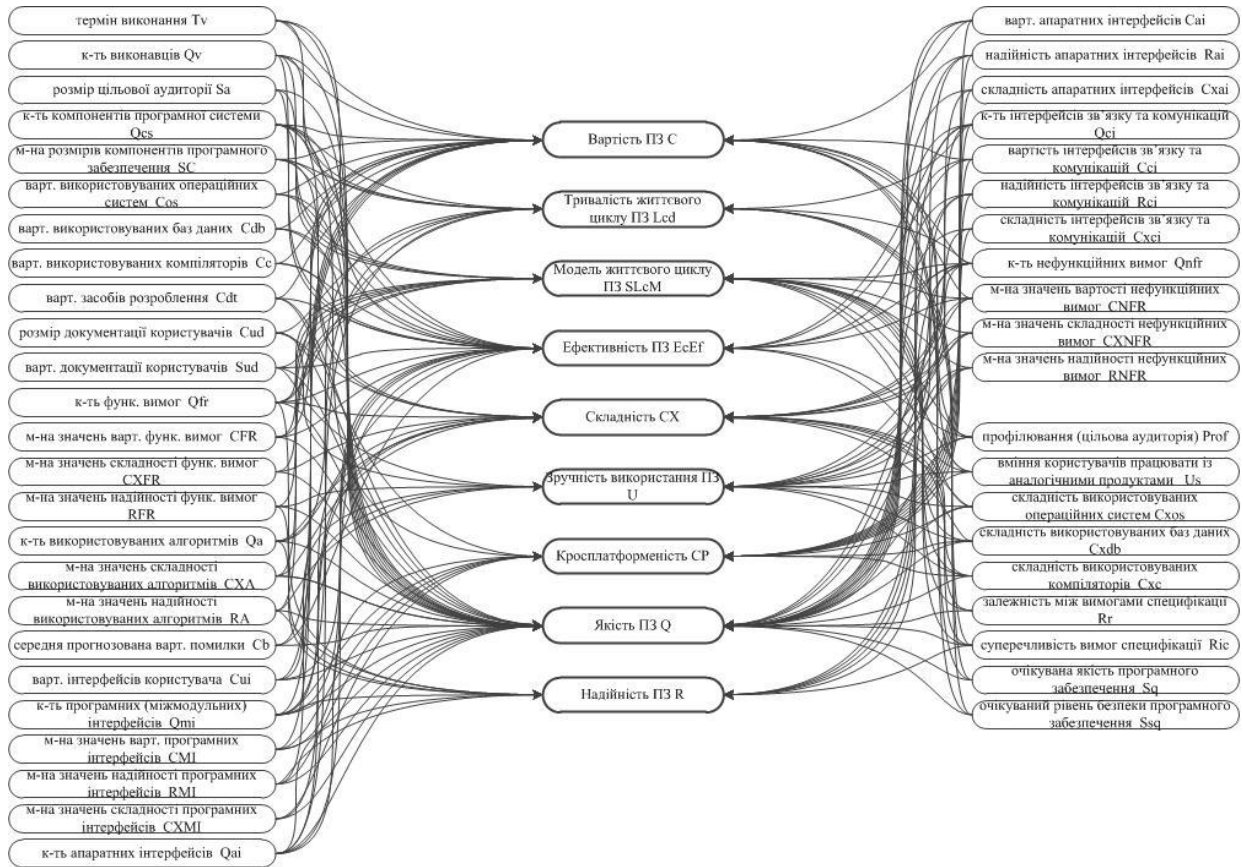


Рис.1. Вплив КП та ЕКП специфікації на характеристики ПЗ

Отже, для розв'язку задачі прогнозування характеристик ПЗ на основі кількісних та експертних кількісних показників специфікації вимог до ПЗ потрібно розв'язати наступну систему рівнянь:

$$\begin{cases}
 C = f(X_C) \\
 LcD = \varphi(X_{LcD}) \\
 SLcM = \phi(X_{SLcM}) \\
 EcEf = \gamma(X_{EcEf}) \\
 CX = \sigma(X_{CX}) \\
 U = \zeta(X_U) \\
 CP = \xi(X_{CP}) \\
 Q = \psi(X_Q) \\
 R = \omega(X_R)
 \end{cases}, \quad (1)$$

де  $X_C = \{x_{C_1}, \dots, x_{C_n}\}$  - множина кількісних та експертних кількісних показників специфікації, які впливають на оцінку вартості ПЗ;  $n$  - кількість КП та ЕКП, які впливають на оцінку вартості ПЗ (згідно вищенаведених моделей характеристик ПЗ  $n = 23$ );  $X_{LcD} = \{x_{LcD_1}, \dots, x_{LcD_m}\}$  - множина КП та ЕКП специфікації, які впливають на оцінку тривалості життєвого циклу ПЗ;  $m$  - кількість КП та ЕКП, які впливають на оцінку тривалості життєвого циклу ПЗ ( $m = 14$ );  $X_{SLcM} = \{x_{SLcM_1}, \dots, x_{SLcM_k}\}$  - множина КП та ЕКП специфікації, які впливають на вибір моделі життєвого циклу ПЗ;  $k$  - кількість КП та ЕКП, які впливають на вибір моделі життєвого циклу ПЗ ( $k = 20$ );  $X_{EcEf} = \{x_{EcEf_1}, \dots, x_{EcEf_l}\}$  - множина КП та ЕКП специфікації, які впливають на оцінку ефективності ПЗ;  $l$  - кількість КП та ЕКП, які впливають на оцінку ефективності ПЗ ( $l = 25$ );  $X_{CX} = \{x_{CX_1}, \dots, x_{CX_s}\}$  - множина КП та ЕКП специфікації, які впливають на

оцінку складності ПЗ;  $s$  - кількість КП та ЕКП, які впливають на оцінку складності ПЗ ( $s=20$ );  $X_U = \{x_{U_1}, \dots, x_{U_s}\}$  - множина КП та ЕКП специфікації, які впливають на оцінку зручності використання ПЗ;  $t$  - кількість КП та ЕКП, які впливають на оцінку зручності використання ПЗ ( $t=17$ );  $X_{CP} = \{x_{CP_1}, \dots, x_{CP_t}\}$  - множина КП та ЕКП специфікації, які впливають на оцінку кросплатформеності ПЗ;  $p$  - кількість КП та ЕКП, які впливають на оцінку кросплатформеності ПЗ ( $p=17$ );  $X_Q = \{x_{Q_1}, \dots, x_{Q_a}\}$  - множина КП та ЕКП специфікації, які впливають на оцінку якості ПЗ;  $a$  - кількість КП та ЕКП, які впливають на оцінку якості ПЗ ( $a=45$ );  $X_R = \{x_{R_1}, \dots, x_{R_b}\}$  - множина КП та ЕКП специфікації, які впливають на оцінку надійності ПЗ;  $b$  - кількість КП та ЕКП, які впливають на оцінку надійності ПЗ ( $b=13$ ).

Побудувати багатовимірне відображення  $X \rightarrow Y$  означає представити його за допомогою математичних операцій над не більш, ніж двома змінними. Теорема Хехт-Нільсена в неконструктивній формі доводить можливість розв'язку задачі представлення функції довільного вигляду на нейронній мережі і вказує для кожної задачі мінімальні кількості нейронів мережі, необхідних для її розв'язку. З теореми Хехт-Нільсена слідує можливість представлення будь-якої багатовимірної функції декількох змінних за допомогою нейронної мережі фіксованого розміру. Згідно теореми про повноту, будь-яка неперервна функція на замкненому обмеженому просторі може бути рівномірно наближена функціями, які обчислюються нейронними мережами, якщо функція активації нейрона двічі неперервно диференційована і неперервна. Таким чином, штучні нейронні мережі (ШНМ) є універсальними структурами, що дозволяють узагальнити інформацію та виявити залежності між вхідними і результуючими даними, отже, для реалізації вищезазначеної системи рівнянь (1) буде використовуватись саме ШНМ. Отже, саме ШНМ опрацьовуватиме множини КП та ЕКП специфікації, здійснюватиме апроксимацію показників та надаватиме прогнозовані оцінки характеристик програмного забезпечення.

Концепцію прогнозування основних характеристик ПЗ на основі аналізу специфікації вимог до ПЗ із використанням штучної нейронної мережі представлено на рис.2.

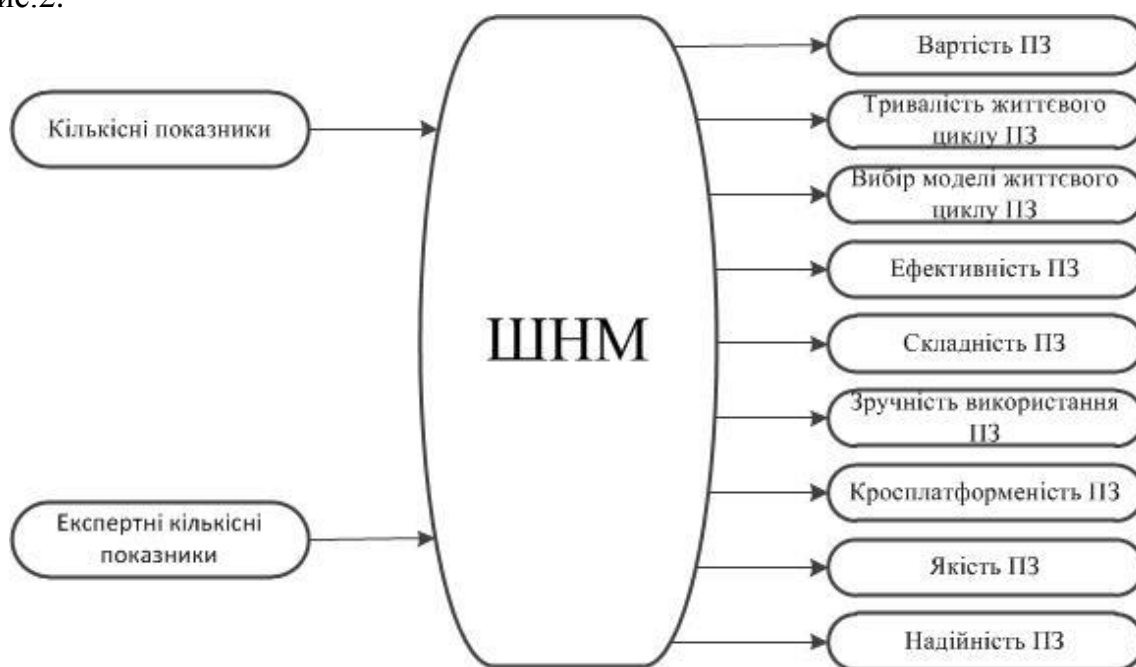


Рис.2. Концепція прогнозування характеристик ПЗ на основі аналізу специфікацій із використанням штучної нейронної мережі

Використовуючи запропоновану концепцію прогнозування характеристик ПЗ на основі аналізу специфікацій, розроблено інтелектуальну модель процесу прогнозування характеристик ПЗ, в основі якої лежить штучні нейронні мережа (ШНМ). Вибір апарату ШНМ мотивований тим, що штучні нейронні мережі за рахунок можливості апроксимації нелінійних функцій дають можливість враховувати важливість (ваги) кожного показника специфікації, а також взаємний вплив показників специфікації в межах кожної характеристики ПЗ. Важкоформалізованою задачею прогнозування характеристик ПЗ є визначення ваг та взаємовпливу кількісних та експертних кількісних показників специфікації в межах кожної характеристики ПЗ. Ця задача може бути вирішена за допомогою використання навченої ШНМ.

Перед розробленням нейромережної моделі прогнозування характеристик ПЗ на основі аналізу специфікацій слід виконати вибір архітектури ШНМ, враховуючи запропоновану концепцію прогнозування характеристик ПЗ. Для цього дослідимо найбільш поширені архітектури ШНМ та їх особливості:

1) багатошаровий перцептрон – найпростіша і найбільш досліджена структура, якої достатньо для задачі прогнозування характеристик ПЗ на основі аналізу показників специфікацій;

2) радіальні базисні мережі (РБМ) вимагають великої кількості навчальних векторів та зазнають значних труднощів, коли є велика кількість входів. Для задачі аналізу показників специфікацій та прогнозування характеристик ПЗ важко зібрати велику кількість навчальної інформації через відсутність великої кількості доступних специфікацій; крім того, має місце велика кількість входів;

3) мережі GRNN (регресійні) – один з видів РБМ, використовуються для аналізу числових рядів. В задачі аналізу показників специфікації та прогнозування характеристик ПЗ зв'язків, аналогічних числовому ряду, немає;

4) мережі PNN (імовірнісні) – один з видів РБМ, призначені для розв'язання імовірнісних задач, зокрема, задач класифікації – встановлення приналежності об'єкту до певного класу. В задачі прогнозування характеристик ПЗ немає потреби в класифікації;

5) карта Кохонена – призначена для кластеризації даних. Ця задача під час аналізу показників специфікації та прогнозування характеристик ПЗ не вирішується;

6) мережі для класифікації входних векторів (LVQ) – використовуються для кластеризації і класифікації. Ці задачі під час прогнозування характеристик ПЗ не розглядаються;

7) мережі Елмана та Хопфілда – це мережі з динамічними оберненими зв'язками, орієнтовані на опрацювання динамічних моделей, що враховують передісторію процесів, які спостерігаються. Задача прогнозування характеристик ПЗ не містить обернених зв'язків і передісторії.

Аналіз відомих архітектур ШНМ показав, що для вирішення задачі аналізу показників специфікацій та прогнозування характеристик ПЗ найбільше підходить багатошаровий перцептрон. При використанні нейромережі іншого типу для розв'язання цієї задачі її природа буде штучно спотворюватись, в результаті чого результати роботи ШНМ будуть невідповідними. У багатошаровому перцептроні нейрони містяться у шарах, що розташовані один за одним, і взаємозв'язки між нейронами шарів передбачені тільки у напрямку від входу до виходу ШНМ.

Відобразимо зазначений підхід узагальненою складною ШНМ архітектури багатошарового перцептрона типу MLP (multi-layer-perceptron). Структура цієї ШНМ представлена на рис.3.

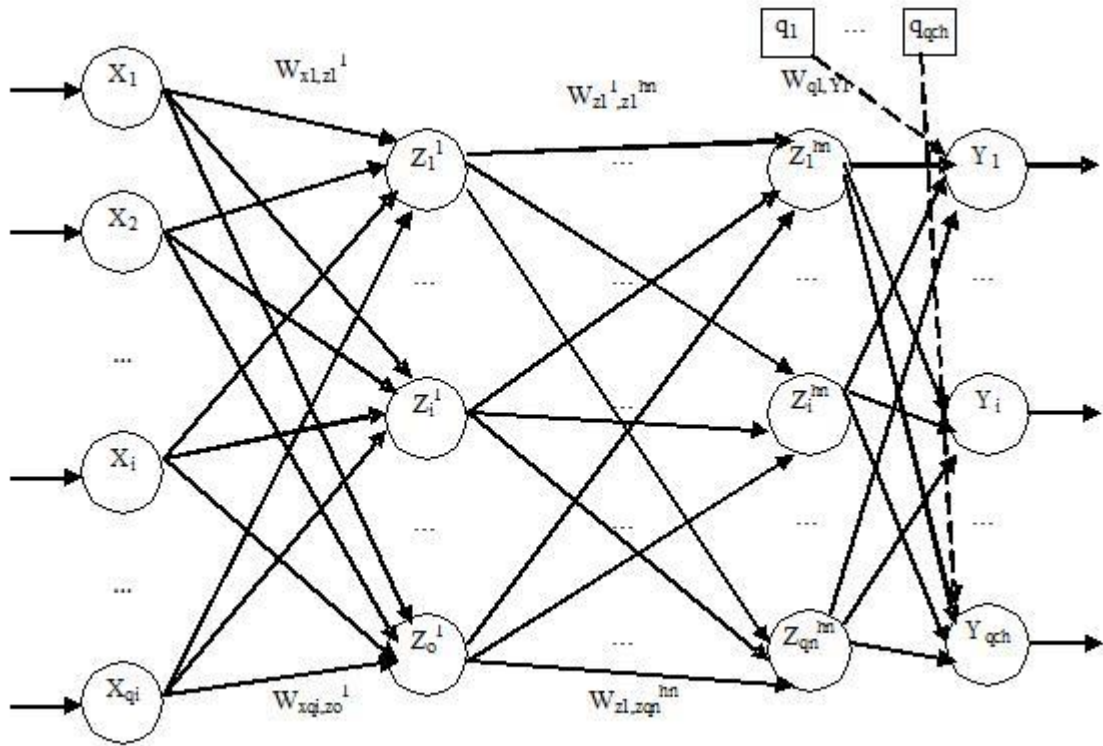


Рис.3. Нейромережна модель прогнозування характеристик ПЗ на основі аналізу специфікацій

Нейрони ( $i$  відповідно входи) рецепторного шару визначено множиною  $\overline{X} = \{x_i\}$ ,  $i = (\overline{1, qi})$ , де  $x_i$  -  $i$ -й нейрон (вхід) рецепторного шару,  $qi$  - кількість входів (кількість КП та ЕКП).

Нейрони першого прихованого шару прямонаправленого персептрона визначено множиною  $\overline{Z}^1 = \{z_i^1\}$ ,  $i = (\overline{1, o})$ , де  $z_i^1$  -  $i$ -й нейрон шару  $z^1$ ,  $o$  - кількість нейронів першого прихованого шару. Рецепторні входи цього шару визначено множиною:  $\overline{X}^{z^1} = \{x^{z_i^1}\}$ ,  $i = (\overline{1, o})$ , де  $x^{z_i^1}$  -  $i$ -й вхід, пов'язаний з  $z_i^1$ -м нейроном. Нейрони другого прихованого шару прямонаправленого персептрона визначено множиною  $\overline{Z}^2 = \{z_i^2\}$ ,  $i = (\overline{1, qn2})$ , де  $z_i^2$  -  $i$ -й нейрон шару  $z^2$ ,  $qn2$  - кількість нейронів другого прихованого шару. Рецепторні входи цього шару визначено множиною:  $\overline{X}^{z^2} = \{x^{z_i^2}\}$ ,  $i = (\overline{1, qn2})$ , де  $x^{z_i^2}$  -  $i$ -й вхід, пов'язаний з  $z_i^2$ -м нейроном. Нейрони  $hn$ -го прихованого шару прямонаправленого персептрона визначено множиною  $\overline{Z}^h = \{z_i^{hn}\}$ ,  $i = (\overline{1, qn})$ , де  $z_i^{hn}$  -  $i$ -й нейрон шару  $z^{hn}$ ,  $qn$  - кількість нейронів  $hn$ -го прихованого шару. Рецепторні входи цього шару визначено множиною  $\overline{Z}^{z^h} = \{z_i^{z^{hn-1}}\}$ ,  $i = (\overline{1, qnm})$ , де  $z_i^{z^{hn-1}}$  -  $i$ -й нейрон попереднього ( $hn-1$ )-го шару, що пов'язаний з  $z_i^{hn-1}$ -м нейроном,  $qnm$  - кількість нейронів ( $hn-1$ )-го прихованого шару.

Вектор порогових величин зміщень множини нейронних елементів визначено як:  $\overline{q} = \{q_i\}$ ,  $i = (\overline{1, qch})$ , де  $q_i$  - зміщення  $i$ -го нейронного елемента.

Ваги зв'язків відображено ваговими матрицями:  
 $\overline{W}_{x,z^1} = \langle w_{x_i, z_j^1} \rangle$ ,  $i = (\overline{1, qi})$ ,  $j = (\overline{1, o})$ , де  $w_{x_i, z_j^1}$  - ваговий коефіцієнт зв'язку між  $x_i$ -м входом та  $z_j^1$ -м нейроном першого прихованого шару;  
 $\overline{W}_{z^1, z^2} = \langle w_{z_i^1, z_j^2} \rangle$ ,  $i = (\overline{1, o})$ ,  $j = (\overline{1, qn2})$ , де  $w_{z_i^1, z_j^2}$  - ваговий коефіцієнт зв'язку між  $z_i^1$ -м

нейроном 1-го прихованого шару та  $z_j^2$ -м нейроном 2-го прихованого шару;  
 $\overline{W_{z^1, z^{hn}}} = \langle w_{z_i^1, z_j^{hn}} \rangle$ ,  $i = (\overline{1, o}), j = (\overline{1, qn})$ , де  $w_{z_i^1, z_j^{hn}}$  - ваговий коефіцієнт зв'язку між  $z_i^1$ -м нейроном 1-го прихованого шару та  $z_j^{hn}$ -м нейроном  $hn$ -го прихованого шару;  
 $\overline{W_{z^{hn-1}, z^{hn}}} = \langle w_{z_i^{hn-1}, z_j^{hn}} \rangle$ ,  $i = (\overline{1, qnm}), j = (\overline{1, qn})$ , де  $w_{z_i^{hn-1}, z_j^{hn}}$  - ваговий коефіцієнт зв'язку між  $z_i^{hn-1}$ -м нейроном  $(hn - 1)$ -го прихованого шару та  $z_j^{hn}$ -м нейроном  $hn$ -го прихованого шару;  
 $\overline{W_{q, Y}} = \langle w_{q_i, Y_i} \rangle$ ,  $i = (\overline{1, qch})$ , де  $w_{q_i, Y_i}$  - ваговий коефіцієнт зв'язку між  $q_i$ -м зміщенням та  $Y_i$ -м нейроном вихідного шару.

Множина вихідних функціоналів ефекторного шару ШНМ становить  $\overline{Y} = \{Y_{ki}\}$ ,  $ki = (\overline{1, qch})$ , де  $Y_{ki}$  -  $ki$ -й функціонал ефекторного шару багат шарового перцептронну,  $qch$  - кількість виходів ШНМ (кількість прогнозованих характеристик ПЗ).

Формула для визначення  $ki$ -го функціоналу ефекторного шару ШНМ  $Y_{ki}$  має вигляд:

$$Y_{ki} = fs \left( fa(z^1) \cdot \left( \sum_{i=1}^o \sum_{j=1}^{qn2} (z_i^1 \cdot w_{z_i^1, z_j^2}) + \dots + \sum_{i=1}^o \sum_{j=1}^{qn} (z_i^1 \cdot w_{z_i^1, z_j^{hn}}) \right) + \dots + fa(z^{hn-1}) \cdot \left( \sum_{i=1}^{qnm} \sum_{j=1}^{qn} (z_i^{hn-1} \cdot w_{z_i^{hn-1}, z_j^{hn}}) \right) + \sum_{i=1}^{qki} (x_i \cdot w_{x_i, z_i}) \right) - q_{ki} \cdot w_{q_{ki}, Y_{ki}} \quad (2)$$

де  $fa(z^1)$  - активаційна функція нейронів 1-го прихованого шару,  $fa(z^{hn-1})$  - активаційна функція нейронів  $(hn - 1)$ -го прихованого шару,  $fs$  - активаційна функція нейронів ефекторного (вихідного) шару ШНМ,  $qki$  - кількість КП та ЕКП специфікації, які впливають на  $ki$ -й вихід ШНМ (на  $ki$ -у прогнозовану характеристику ПЗ),  $x_i$  -  $i$ -й елемент підмножини КП та ЕКП, які впливають на  $ki$ -й вихід ШНМ.

Активаційною функцією нейронів прихованих (асоціативних) шарів є функція гіперболічного тангенсу. Активаційною функцією нейронів ефекторного шару є лінійна функція. Результати лінійної активаційної функції нейронів ефекторних шарів лежать в інтервалі  $[-1; 1]$ .

Запропонована концепція прогнозування характеристик ПЗ на основі аналізу специфікацій дає можливість розробити інтелектуальну модель процесу прогнозування характеристик ПЗ, в основі якої лежить штучна нейронна мережа (ШНМ).

Розроблена нейромережна модель прогнозування характеристик ПЗ на основі аналізу специфікацій (НМПХ) відрізняється від існуючих тим, що дає можливість враховувати ваги та взаємовпливи показників специфікації в межах кожної характеристики ПЗ. Вихідні функціонали ШНМ, що відповідають прогнозованим значенням характеристик ПЗ, дають можливість оцінити сумарний вплив КП та ЕКП специфікації на характеристики розроблюваного за специфікацією ПЗ і зробити висновки щодо обґрунтованого вибору програмного проекту.