

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КІЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»**

Факультет прикладної математики

Кафедра прикладної математики

«До захисту допущено»

Завідувач кафедри  
\_\_\_\_\_ О.Р. Чертов

“ \_\_\_\_ ” 2015 р.

**Дипломна робота**

на здобуття ступеня бакалавра

зі спеціальності 6.040301 «Прикладна математика»

на тему: Математичне моделювання та оптимізація стратегій прийняття  
рішень в терапії пневмонії

Виконав: студент 4 курсу, групи КМ-12

Данилевський Володимир Ігорович

Керівник старший викладач Любашенко Н.Д.

Консультант з наукової частини доцент Соловйов С.О.

Консультант з нормконтролю ст. викладач Мальчиков В.В.

Рецензент доцент Галицька І.Є.

Засвідчую, що у цій дипломній  
роботі немає запозичень з праць  
інших авторів без відповідних  
посилань.

Студент \_\_\_\_\_

Київ – 2015 року

**Національний технічний університет України  
«Київський політехнічний інститут»**

Факультет прикладної математики  
Кафедра прикладної математики  
Рівень вищої освіти – перший (бакалаврський)  
Спеціальність 6.040301 «Прикладна математика»

ЗАТВЕРДЖУЮ  
Завідувач кафедри  
\_\_\_\_\_ О. Р. Чертов  
«\_\_\_\_ » \_\_\_\_\_ 2015р



1. Тема роботи «Математичне моделювання та оптимізація стратегій прийняття рішень в терапії пневмонії»,  
керівник роботи Любашенко Наталія Дмитрівна, старший викладач,  
 затверджені наказом по університету від “19” травня 2015 року №1039-С.
2. Строк подання студентом роботи: “12” червня 2015 р.
3. Вихідні дані для дипломної роботи:
  - дані клінічних спостережень за період 2013 рр, та даних медичних карток хворих на негоспітальну пневмонію Чернігівського військового госпіталю.
4. Перелік задач, які потрібно вирішити:
  - вивчити літературні джерела за тематикою дослідження, в тому числі методики побудови математичних моделей в біонформатиці;
  - провести аналіз математичних методів прогнозування в біонформатиці;
  - обґрунтувати вибір методу моделювання;

- обрати необхідні ключові параметри для включення в математичну модель
- знаходження оптимального шляху терапії хворого
- озробити математичну модель для прогнозування середньої тривалості госпіталізації хворого та фармакоекономічну оцінку препаратів ;
- створити програмний продукт для обробки цих даних по цій моделі;
- оцінити адекватність висновків

**5. Перелік обов'язкового ілюстративного матеріалу:**

- теоретичні аспекти побудови математичної моделі для прогнозування у біоінформації;
- аспекти побудови програмного продукту;
- результат роботи програмного продукту.

**6. Консультанти розділів проекту(роботи):**

Питання	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Ключові параметри негоспітальної пневмонії	Соловйов С.О. доцент.		
Підходи прогнозування у біоінформації	Соловйов С.О. доцент.		
Теорія фармакоекономічної оцінки	Соловйов С.О. доцент.		

7. Дата видачі завдання: “31” жовтня 2014 р.

### КАЛЕНДАРНИЙ ПЛАН

№ з/ п	Назва етапів виконання дипломної роботи	Срок виконання етапів	Приміка
1.	Вивчення літератури за тематикою роботи та збір даних	15.11.2014	
2.	Проведення порівняльного аналізу математичних методів  У біонформатиці	10.12.2014	
3.	Підготовка матеріалів першого розділу дипломної роботи	30.12.2014	
4.	Розробка математичної моделі для прогнозування	15.01.2015	
5.	Підготовка матеріалів другого розділу дипломної роботи	10.02.2015	
6.	Підготовка тез доповіді на конференції «Прикладна математика та комп’ютинг»	20.02.2015	
7.	Підготовка матеріалів третього розділу дипломної роботи	10.03.2015	
8.	Підготовка матеріалів четвертого розділу дипломної роботи	10.04.2015	
9.	Підготовка графічної частини дипломної роботи	20.05.2015	
10	Оформлення дипломної роботи	27.05.2015	
-			

Студент \_\_\_\_\_ Данилевський В.І.

Керівник роботи \_\_\_\_\_ Любашенко Н.Д.

## АНОТАЦІЯ

Дипломна робота присвячена розробці математичної моделі та програмного продукту для знаходження оптимальної схеми терапії пацієнта на негоспітальну пневмонію та розрахунку фармакоекономічної ефективності препаратів для існуючих терапій даної нозологічної одиниці.

Дипломна робота містить вступ, перелік умовних скорочень, дев'ять розділів, висновки, перелік посилань, ілюстративні матеріали та додатки.

В роботі виконано порівняльний аналіз математичних методів математичного моделювання у біоінформатиці, зокрема, розглянуто та проаналізовано підхід з побудовою динамічної баєсової мережі та дерева можливостей для оцінки фармакоекономічної ефективності обраної терапії. Розрахунки виконуються для заданих часових проміжків, а результат відображає стан хворого у певний проміжок часу. Вхідними даними є база існуючих клінічних рішень та стан даного хворого, а результатом є оптимальна схема лікування хворого та економічна ефективність обраної терапії.

Робота також має у собі програмний продукт який може доповнювати, корегувати існуючі дані та на основі цих даних отримувати прогноз оптимальної терапії для нових вхідних даних.

Дана робота може бути використана у закладах охорони здоров'я для покращення обслуговування хворих, зменшення навантаження на співробітників медичних закладів.

Перелік ключових слів: Біоінформатика, Баєсова мережа, Пневмонія, Фармакоекономіка, Математична модель.

## ABSTRACT

This diploma is devoted to the development of mathematical models and software for finding the optimal way of therapy patients with community acquired pneumonia and calculation farmoekonomics efficacy of chosen therapies.

This diploma contains introduction, nine main chapters, conclusion, illustrated materials, linked literature, appendix.

In this work was made the comparative analysis of mathematical methods for modeling in bioinformatics. In particular, reviewed and analyzed the approach of building dynamic bayesian networks and tree of opportunities for pharmacoconomics evaluation of efficiency of the chosen therapy. Calculations are performed for the specified period, the result represent the condition of the patient at determinate time. Input data is the existing date in base of clinical, and the result is the optimal treatment regimen for patient and cost-effectiveness of this therapy.

Work also include a software product that can add, correct existing data and based on these data, produce optimal therapy for new input.

This diploma can be used for medical institutes for decreasing operation time per patients and make easiest work for personal.

**Key words:** Bioinformatics, Bayesian network, Pneumonia, Farmoekonomics, Mathematical model.

## ЗМІСТ

СПИСОК ТЕРМІНІВ, СКОРОЧЕНЬ ТА ПОЗНАЧЕНЬ .....	9
ВСТУП .....	10
1 МЕТА ДОСЛІДЖЕННЯ ТА ПОСТАНОВКА ЗАДАЧІ .....	12
2 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ .....	13
2.1 Медична діагностика .....	16
2.2 Медична прогностика .....	20
2.3 Висновки .....	22
3 ІНДИВІДУАЛЬНЕ ЗАВДАННЯ .....	23
3.1 Висновки .....	26
4 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ .....	27
4.1. Огляд літературних джерел .....	28
4.2. Огляд методів вирішення поставленої задачі .....	29
4.3. Вибір вирішення поставленої задачі .....	31
4.4. Виділення параметрів з предметної області .....	32
4.5 Висновки .....	34
5 ВИКОРИСТАННЯ БАЄСОВОЇ МЕРЕЖІ .....	35
5.1. Побудова баєсової мережі .....	35
5.2. Опис алгоритму обчислення ймовірностей у динамічних баєсових мережах .....	36
5.3 Висновок .....	38
6 ФАРМАКОЕКОНОМІЧНА ЕФЕКТИВНІСТЬ ТЕРАПІЇ .....	39
6.1. Фармацоекономічна ефективність .....	39
6.2. Фармацоекономічна модель у даній задачі .....	39
6.3 Висновок .....	41
7 ОБРАННЯ ШЛЯХУ ПРОГРАМНОЇ РЕАЛІЗАЦІЇ .....	42
7.1 Висновки .....	43
8 ПРОГРАМНА РЕАЛІЗАЦІЯ .....	44
8.1 Робота у програмі .....	46

8.2 Висновки .....	49
9 ПРОГНОЗУВАННЯ .....	50
ВИСНОВКИ.....	51
СПИСОК ЛІТЕРАТУРИ.....	52



## СПИСОК ТЕРМІНІВ, СКОРОЧЕНЬ ТА ПОЗНАЧЕНЬ

$v$  – символ для позначення набору випадкових величин

$\text{dom}(x)$  – область можливих значень

$x = \times$  - Змінна  $x$  у стані  $\times$

$p(x = \text{tr})$  – Вірогідність того, що зміна  $x$  у стані *true*

$p(x = \text{fa})$  – Вірогідність того, що зміна  $x$  стані *false*

$p(x, y)$  – Вірогідність одночасного випадку  $x$  та  $y$

$p(x \cap y)$  – Вірогідність одночасного випадку  $x$  та  $y$

$p(x \cup y)$  – Вірогідність хоча б одного з випадків  $x$  або  $y$

$pa(x)$  – Батьківський елемент вершини  $x$

$ch(x)$  – Дочірній елемент вершини  $x$

$ne(x)$  – Сусідній елемент до вершини  $x$

БМ – Баєсова мережа

ДБМ – Динаміна баєсова мережа

НП – Негоспітальна пневмонія

## ВСТУП

Прогнозування динаміки чисельності та структури населення, складу сімей, режимів відтворення населення, різних демографічних процесів складає значну частину діяльності міжнародних, державних і неурядових організацій, закладів та наукових інститутів. Провідні демографи, економісти, математики здійснили чимало глибоких, різноаспектних досліджень проблем моделювання відтворення населення (приросту населення).

Прогнозування стану пацієнта та підбір шляху терапії хворого є розповсюдженою задачею у сфері біоінформатики. Також у даний час лікувально-профілактичні заклади проходять стадію переходу від паперового документообігу до автоматизованого обліку даних, зокрема існують комп'ютеризовані системи зберігання даних пацієнтів. Тому отримання великого масиву даних, а саме: даних щодо захворювання, клініко-лабораторних даних хворого, обраної схеми терапії, вимагає зручного аналізу таких даних з використанням системи підтримки прийняття рішень. Головною проблемою при створенні такої системи є неповнота даних, якими вона оперує. Таким чином, необхідне використання відповідного математичного апарату, який зможе опрацьовувати такі дані. Одним з таких підходів є використання динамічної баєсової мережі (ДБМ)[1], яка дозволяє працювати в умовах неповноти вхідних даних.

Дана тематика це така модель, що буде залишатись актуальною, оскільки автоматизація медичних закладів, зменшення навантаження на працівників медичних закладів є актуальним напрямом оптимізації роботи. А теоретичне зменшення терміну госпіталізації завдяки більш точному підбору терапії. Також це дозволяє оцінити ефективність існуючих медикаментозних препаратів, та проаналізувати їх фармакологічну цінність.

Отже, створення подібної ІС дозволить зменшити витрати на госпіталізацію хворих у подальшому, полегшити роботу медичному персоналу, надати інструмент для оцінки ефективності лікарських засобів для лікування НП.

Об'єктом дослідження було обрано НП тому що це одне з найбільш розповсюджених захворювань дихальної системи, тому ми маємо достатньо даних для аналізу та перевірки на адекватність отриманої моделі.

У сучасній біоінформатиці застосовують методи математичного моделювання для прогнозування того чи іншого параметру, у рамках цієї дипломної роботи - це буде прогнозування стану хворого, методу його лікування та ефективності лікування вираженої через термін лікування, перебування хворим кількості ліжко-днів в стаціонарі та ефективності проведеної терапії.

Для подібних задач існують такі основні підходи:

1. Нейроні мережі
2. Баєсові мережі

Також до завдання було включено завдання знаходження фармакоекономічної ефективності медичного препарату, для подібної оцінки використовують дерево прийняття рішення.

У ході аналізу математичних методів було обрано метод на основі БМ, тому що цей метод надає можливість працювати з неповними даними, а також має перевагу в подальшому легко отримати новий або змінити старий параметр у мережі. Також його легко перевірити на адекватність.

## 1 МЕТА ДОСЛІДЖЕННЯ ТА ПОСТАНОВКА ЗАДАЧІ

Метою дослідження є математичне моделювання для визначення оптимального шляху терапії хворого на НП та оцінки фармакоекономічної ефективності даного шляху.

Науково-практична задача, що розв'язується в даній дипломній роботі, включає наступні завдання:

- а) Провести аналіз математичних методів прогнозування в біонформатиці;
- б) Обґрунтувати вибір методу моделювання;
- в) Обрати необхідні ключові параметри для включення в математичну модель
- г) Розробити математичну модель для прогнозування середньої тривалості госпіталізації хворого та фармакоекономічну оцінку використовуваних препаратів
- д) Знаходження оптимального шляху терапії хворого
- е) Створити програмний продукт для обробки цих даних по цій моделі
- ж) Оцінити результати з існуючими практичними даними

## 2 ОГЛЯД ПРЕДМЕТНОЇ ОБЛАСТІ

В умовах невизначеності можливий розвиток небажаних подій, що несуть за собою економічний, моральний або інший збиток людині та суспільству в цілому. Такі події характеризуються ризиком їх реалізації. Для значної кількості областей знань загальним є уявлення про ризик як про характеристику несприятливих наслідків невизначеності ситуації. У таких областях для зменшення наслідків реалізації ризику використовуються методи ризик-менеджменту, які включають визначення та вимірювання ризиків, оцінки ризиків та управління ними.

Оцінка медичних ризиків безпосередньо використовується при страхуванні життя і здоров'я людини. Згідно, для цілей страхування виділяють декілька рівнів ризику, пов'язаного з завданням шкоди здоров'ю суб'єкта:

- 1) Ризик розвитку хвороби або фізичного травматизму є ризик фізіологічних порушень, які загрожують нормальному функціонуванню організму. Наслідки реалізації цього ризику мають безпосередній зв'язок з нанесенням економічного збитку як суб'єкту, так і (в окремих випадках) державі.
- 2) Ризик необхідності відновлення здоров'я (лікування). Цей рівень ризику має певний економічний сенс - витрати на ресурси, що будуть затрачені на лікування.
- 3) Ризик порушень соціально-економічної сторони життя у зв'язку з втратою здоров'я. Збиток при реалізації цього рівня ризику має як соціальну, так і економічну складову.
- 4) Ризик необхідності підтримки прийнятних умов життя при незворотному і значимому збитку для здоров'я. Так само, як і в разі реалізації ризиків другого рівня, збиток від реалізації цього рівня ризику може бути

виражений у грошовій формі, а саме у вигляді витрат на необхідні ресурси.

Ризик в медицині пов'язаний з прийняттям рішень в умовах невизначеності, яка в більшості випадків пов'язана зі значним впливом людського фактора на процес лікування. Лікар змушений вибирати як серед деякої множини альтернативних діагнозів (в силу неповноти і неточності інформації про симптомах і результатах лабораторних обстежень), так і серед деякого безлічі альтернативних підходів до діагностики та лікування захворювання, зважуючи при цьому можливий розвиток побічних ефектів у пацієнта при проведенні медичного втручання, і можливі вигоди для нього. Таким чином, завжди є ризик заподіяння шкоди пацієнту під час його лікування. Згідно з проведеними дослідженнями, в США при роботі з 45% пацієнтів була допущена та чи інша медична помилка, в 17% випадків така помилка призвела до більш тривалого перебування в лікарні або погіршення стану здоров'я пацієнта [2]. Виникнення небажаних наслідків при медичному втручанні тісно пов'язане з багатьма факторами, серед яких проблеми нагляду, перевантаження на робочому місці, недоліки в навчанні фахівців, проблеми при роботі в команді. Задача оцінки рівня безпеки лікування пацієнта і ризику виникнення несприятливих для нього наслідків вимагає системного підходу в силу значної складності процесів. Медичні системи, з точки зору оцінки ризику нанесення шкоди пацієнту, можуть розглядатися як соціо-технічні системи, велике значення в яких має людський фактор. Складна медична система може бути представлена як адаптивна система, в якій групи людей кооперуються з метою турботи про пацієнта і стикаються з безліччю непередбачених обставин, і ключовий характеристики.

Виникнення лікарських помилок веде до додаткових витрат системи охорони здоров'я і повинно враховуватися при управлінні ризиками установи системи охорони здоров'я.

Сучасні медичні системи повинні задовольняти вимогам безпеки пацієнтів та надання дієвої та ефективної допомоги. В силу складності та значною невизначеності медичної системи, недостатньо застосовувати описані моделі для оцінки ризику та методичні вказівки для збільшення ступеня безпеки турботи про пацієнта. У наш час пропонується значна кількість різних моделей інформаційних систем, програмних і технічних продуктів (в т.ч. медичного обладнання), які інтегруються в різні частини медичної системи і орієнтовані на поліпшення рівня якості турботи про пацієнтів. Помітна частка зусиль приділено розвитку систем підтримки прийняття рішень в медицині. Кожна нова технологія повинна оцінюватися з точки зору можливостей та ефективності її використання в конкретній медичній системі, що тісно пов'язане з аналізом економічних складових системи. В роботі описана модель, яку можна використати при ретроспективному аналізі ризиків при наданні допомоги пацієнтам.

Специфіка виникнення небажаних наслідків при медичному втручанні в життя пацієнтів знаходить своє відображення в ризик-менеджменті заклади охорони здоров'я. Ризик-менеджмент як процес прийняття рішень, спрямованих на зниження неприємних наслідків реалізації ризикових ситуацій, спрямований на підвищення економічної ефективності діяльності установи.

Управління ризиками медичної системи тісно пов'язане з управлінням безпекою надання медичної допомоги. Виявлення причинно-наслідкових зв'язків, що тягнуть за собою лікарські помилки, може служити основою для поліпшення якості обслуговування. Згідно міжнародним принципам управління ризиками, така система включає в себе етапи вимірювання ризику, оцінки ризику, та управління ризиком.

В рамках цієї роботи розглядається імовірнісний підхід до оцінки ризику. Баєсовські мережі мають ймовірнісну семантику, використовуються для міркувань в умовах невизначеності і все частіше застосовуються в діагностиці захворювань, виборі оптимального курсу

лікування пацієнта, прогнозуванні результату захворювання, побудові моделей захворювань в клінічній епідеміології.

Мета даної роботи полягає в застосування апарату баєсівських мереж для оцінки медичних ризиків та підтримки прийняття рішень в умовах невизначеності.

## 2.1 Медична діагностика

Методи штучного інтелекту знайшли широке застосування в області діагностики захворювань. Інтелектуальні системи в галузі медичної діагностики спрямовані на підтримку прийняття рішення про характер захворювання практикуючим лікарем на підставі результатів різних медичних обстежень і даних про самопочуття пацієнта. Більшість таких систем носять перспективний характер, тобто мають вигляд правил: якщо є такі спостереження, то необхідно робити таке втручання. Однак, як було зазначено, існуючим лікарям доводиться працювати в умовах невизначеності симптомів та результатів клінічних тестів.

Наприклад, недосвідчений лікар здатний коректно поставити діагноз, пов'язаний з захворюванням печінки, менш ніж в 45% випадків. Такий низький відсоток правильної діагностики пов'язаний з тим, що захворювання печінки можна чітко диференціювати лише на певній стадії захворювання. Невірний діагноз тягне як шкоду здоров'ю пацієнта, так і неефективне використання людських, економічних ресурсів. В даний час медична діагностика вимагає врахування значної кількості факторів: генетичні фактори, навколишнє середовище, стиль життя пацієнта. Крім того, галузі біомедицини та охорони здоров'я включають вивчення процесів, які пов'язані з невизначеністю перебігу фізіологічних процесів в організмі людини.

Баєсові мережі дозволяють моделювати знання з невизначеністю. Апарат байєсівської мережі довіри дозволяє створювати комбінацію наявних статистичних даних та характеристи здоров'я пацієнтів на додаток до експертної інформації, яку надають лікарі-фахівці. Крім того, баєсові мережі (порівняно з іншими моделями) дозволяють моделювати можливість виникнення декількох захворювань і елементи БМ мають досить просту інтерпретацію. Велика частина діагностичних моделей підтримки прийняття рішень на основі баєсівських мереж довіри будується на основі експертної інформації, що частково пояснюється особливостями наявної медичної статистики. Особливістю медичної статистики є її часткова коректність і недостатність для автоматичної побудови баєсової мережі. Навіть експертної інформації часто буває недостатньо для побудови повноцінної моделі. Важливою перевагою моделі, спираються на принципи штучного інтелекту, є можливість автоматичного навчання структури моделі, тобто навіть якщо первісна структура моделі була неповною, то мається можливість поліпшити модель за допомогою даних, що надходять до моделі. Крім того, в силу неповноти даних моделі, в ряді випадків застосовні приблизні оцінки результатів імовірнісного виводу в діагностичній баєсовій мережі.

Структура діагностичної баєсової мережі довіри представлена елементами процесу діагностики захворювання і взаємозв'язками між ними. Основою для побудови взаємозв'язків є експертна інформація, наявні діагностичні методи дослідження, спрямовані на виявлення подібних взаємозв'язків між елементами процесу діагностики захворювання, накопичені статистичні дані. Як уже згадувалося, найчастіше структура діагностичної баєсової мережі будується на основі експертної інформації. Кожен вузол являє собою той чи інший фактор, що впливає на виникнення захворювання, і описується випадковим елементом. Випадковий елемент може приймати дискретні значення (чи є пацієнт курцем чи ні) або приймати значення з деякого інтервалу (температура тіла). Апарат баєсової мереж

дозволяє комбінувати випадкові елементи дискретного і безперервного характеру при баєсовому виведенні, проте часто безперервні зміні дискретизуються згідно з наявними медичними знаннями. Таким чином, вузли діагностичної байєсівської мережі описують стан пацієнта. Вузли захворювань можуть представляти собою як бінарні випадкові елементи, що визначають наявність або відсутність захворювання, або ж номінативний випадковий елемент, кожне значення якого відповідає тому або іншому захворюванню. Апріорне розподіл ймовірності на вузлах захворювань визначається показником поширеності захворювання.

В описаній моделі формально найбільш ймовірний діагноз можна визначити як значення множини допустимих діагнозів, максимум ймовірності наявного захворювання при конкретного набору фактів Е, які включають в себе симптоми, результати тестів і інші фактори.

У найпростішому (і найчастіше використовуваному) випадку завдання діагностики захворювання моделюється за допомогою (наївного) баєсова класифікатора. Для побудови моделей баєsovих класифікаторів використовується три типи змінних: спостережувані зміні, відповідні безпосередні дані, змінні хвороб, латентні змінні, не доступні для спостереження, але значення яких може бути важливо при діагностиці. Модель наївного баєсового класифікатора не враховує клас латентних змінних і припускає наявність умовної незалежності спостережуваних змінних один від одного при значенні змінної захворювання (змінної класу). Більш складні моделі баєsovих класифікаторів включають розширені на дерево баєсової мережі. Введення особливих латентних змінних так само служить розширенням моделі наївного баєсового класифікатора, і може бути цікавою з медичної точки зору. Незважаючи на практичну неправдоподібність пропозиції про умовну незалежності факторів захворювання, модель наївного баєсового класифікатора має в ряді випадків таку ж точність прогнозу, що і більш складні моделі. Коректність діагностики (класифікатора) за допомогою інтелектуальної системи

оцінюється трьома показниками: прогнозована цінність позитивного результату, чутливість, специфічність, ROC-аналіз (receiver operating characteristic analysis) дозволяє оцінити якість бінарної класифікації. Крива ROC відображає залежність частки вірних позитивних класифікацій (чутливість) від частки помилкових позитивних класифікацій (одиниця мінус значення показника специфічності) при варіюванні порога вирішального правила. Чисельною характеристикою якості моделі класифікації служить значення площі під кривою ROC, так званий показник AUC (area under the curve). Чим більше площа під кривою ROC, тим вище прогностична точність класифікації.

Емпірично доведено, що інтелектуальні системи діагностики захворювань, засновані на баєсовых мережах довіри, показують кращу діагностичну точність, ніж лікарі-фахівці, або прийняті методики діагностики. Наприклад, інтелектуальна система на основі баєсової мережі для термінової діагностики пневмонії (поза госпітальної пневмонії або CAP community-acquired pneumonia), побудована за статистичними даними, має чутливість 95%, специфічність 96,5%, PPV 26,8% (що означає, що з чотирьох сигналів системи про наявність захворювання, тільки один буде правильним), показник AUC 0,98.

Апарат баєсовых мереж пропонує методи для вибору підходящого медичного обстеження для зменшення невизначеності при діагностиці захворювання у конкретного пацієнта. Такі методи засновані на побудові інформаційних заходів визначеності при постановці діагнозу.

## 2.2 Медична прогностика

Медичний прогноз визначається як прогноз майбутнього курсу і результату процесів, пов'язаних із захворюванням, як у випадку лікування, так і в разі природного перебігу захворювання. Моделі медичної прогностики використовуються як на індивідуальному рівні для підтримки прийняття рішень про майбутній курс лікування, так і на рівні груп пацієнтів для управління ресурсами в охороні здоров'я. Прогностичні баєсові мережі дозволяють обійти труднощі, з якими стикаються традиційні прогностичні моделі. Передбачення в таких моделях зазвичай являє собою однокроковий процес і не передбачає внесення в модель додаткових змінних, які стають відомі з часом. Крім того в загальному випадку не всі змінні, здатні вплинути на результат, включаються в модель, наприклад, якщо два предиктора залежні, то зазвичай включається лише один з них.

Медичний прогноз являє собою міркування в умовах невизначеності. Результат захворювання залежить від послідовності прийнятих рішень про лікування, перебігу хвороби в конкретному випадку і початкового стану пацієнта. формально прогноз визначається як ймовірність за формулою (2.1):

$$Pr(outcome|E, T) \quad (2.1)$$

де Е - доступна інформація про пацієнта (симптоми, результати обстежень, інші ознаки), Т - обрана схема лікування. Результатом медичного втручання може служити як очікувана прогнозована тривалість життя пацієнта, так і інші аспекти якості його життя, як наприклад час госпіталізації. Якість сучасної роботи з пацієнтами залежить від економічного та етичного обґрунтування кожного медичного втручання в життя пацієнта. Принципи

доказової медицини стверджують, що діагностика захворювань і призначення лікування тісно взаємопов'язані і мають здійснюватися з урахуванням медичного прогнозу для пацієнта. Приміром, щоб піти на велике хірургічне втручання, необхідно взяти до уваги якість життя пацієнта, очікувану тривалість його життя, небажані наслідки прийому лікарських засобів, можливість розвитку ускладнень, бюджетні обмеження пацієнта і вартість операції. Прогностичні моделі вимагають обліку значного числа факторів, частина з яких виходить за рамки компетенції лікаря-спеціаліста.

Математичні моделі медичного прогнозування, в тому числі дозволяють класифікувати поточний стан пацієнта по категоріям ризику, пов'язаним з його майбутнім станом. Такі моделі дозволяють чисельно вимірюти тяжкість захворювання і стан здоров'я пацієнта, що використовується в страхуванні життя і здоров'я. Кількісна оцінка ризику альтернативних медичних втручань дозволяють уникнути лікарських помилок і приймати обґрунтовані рішення, як з боку лікаря, так і з боку пацієнта. В якості основи для математичних моделей прогнозування стану здоров'я пацієнтів використовуються теорія прийняття рішень, що є синтезом теорії ймовірностей і теорії вибору, і методи штучного інтелекту. Моделі, що спираються на принципи теорії прийняття рішень, включають в себе дерево прийняття рішень.

Прогностичні баєсові мережі довіри почали розвиватися лише недавно. Баєсові мережі довіри дозволяють представляти як тимчасові, так і позачасові взаємозв'язки між змінними, що роблять вплив на медичний прогноз, що обумовлює можливості використання цієї моделі в медичному прогнозуванні. Відомі приклади побудови прогностичних баєсовых мереж в областях медицини: онкології та інфекційних захворювань.

Завдання вибору курсу лікування тісно пов'язана з визначенням прогнозу для всіх альтернативних варіантів медичного втручання в життя пацієнта. Баєсові мережі довіри не надають апарату для безпосереднього

прийняття рішень про вибір курсу лікування серед деякого набору можливих варіантів, однак використаються як складова частина такого прогностичного міркування, яке спирається на методи теорії прийняття рішень. Так само можливе розширення байєсівської мережі довіри, яку включає в себе інформацію про переваги в рішеннях.

### 2.3 Висновки

Виходячи з аналізу предметної було визначено основні вимоги до прогностичних систем у сфері охорони здоров'я, вимоги до точності результатів подібних систем. Також було встановлено які можливості може надати такі системи, а саме зменшення ризику при госпіталізації, покращення стану пацієнта, зменшення навантаження на медичний персонал.

## З ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

Індивідуальне завдання роботи на здобуття ступеня бакалавра було отримано на минулій виробничій практиці та після уточнення з науковим керівником має наступний вигляд:

### **МОДЕЛЮВАННЯ ТА ОПТИМІЗАЦІЯ СТРАТЕГІЙ ПРИЙНЯТТЯ РІШЕНЬ В ТЕРАПІЇ ПНЕВМОНІЙ**

У цю роботу входять наступний перелік задач котрі повинні бути вирішенні:

- вивчення літературних джерел за тематикою дослідження, в тому числі методики побудови математичних моделей в біонформатиці;
- проведення аналізу математичних методів прогнозування в біонформатиці;
- обґрунтування вибору методу моделювання;
- обрання необхідних ключових параметрів для включення в математичну модель
- знаходження оптимального шляху терапії хворого
- розробка математичної моделі для прогнозування середньої тривалості госпіталізації хворого та фармакоекономічну оцінку препаратів ;
- створення програмного продукту для обробки цих даних по цій моделі;
- оцінити адекватність висновків

Перелік обов'язкового матеріалу:

- теоретичні аспекти побудови математичної моделі для прогнозування у біоінформатиці;
- алгоритм прогнозування;
- аспекти побудови програмного продукту

Об'єктом автоматизації у програмному продукті є медичні картки хворих, які проходять лікування у медичних закладах, отримання діагнозу та схеми терапії що отримують хворі.

Обмін даними концептуально можливо представити наступним шляхом зображенням на Рисунку 3.1:

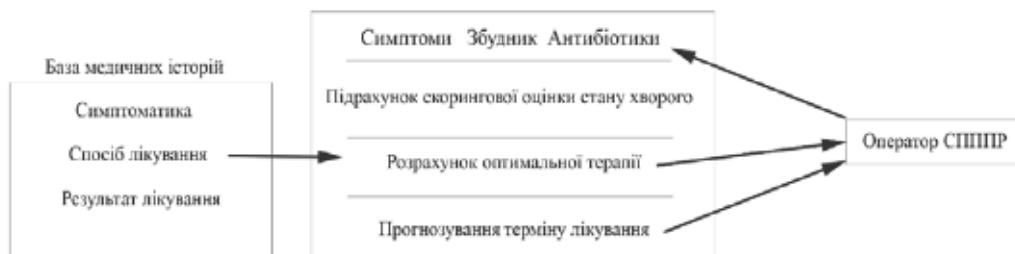


Рисунок 3.1 - Модель обміну даними

Так існує база даних усіх минулих історій хвороб інших пацієнтів, які вміщують у себе симптоматику, спосіб лікування та результат госпіталізації. Оператор від медичного закладу необхідно ввести нові дані до системи, а саме вказати збудника, симптоми, можливе лікування та дані медичної картки хворого (історія хвороб, вік та інше). Система отримавши ці дані має зробити висновки з цих показників, та надати оператору за короткий проміжок часу, не більше хвилини, інформацію про скориговану оцінку стану хворого, розрахувати для хворого оптимальну терапію, котра зменшить час госпіталізації. Також для знайденої схеми терапії визначити прогнозований час госпіталізації.

Програмний продукт має бути реалізований у вигляді тонкого клієнту, обчислення виконуються на стороні сервера, швидкість обчислень не залежить від можливостей користувача. Користувач має можливість увести як пакетно так і через зручний графічний інтерфейс користувача необхідну інформацію.

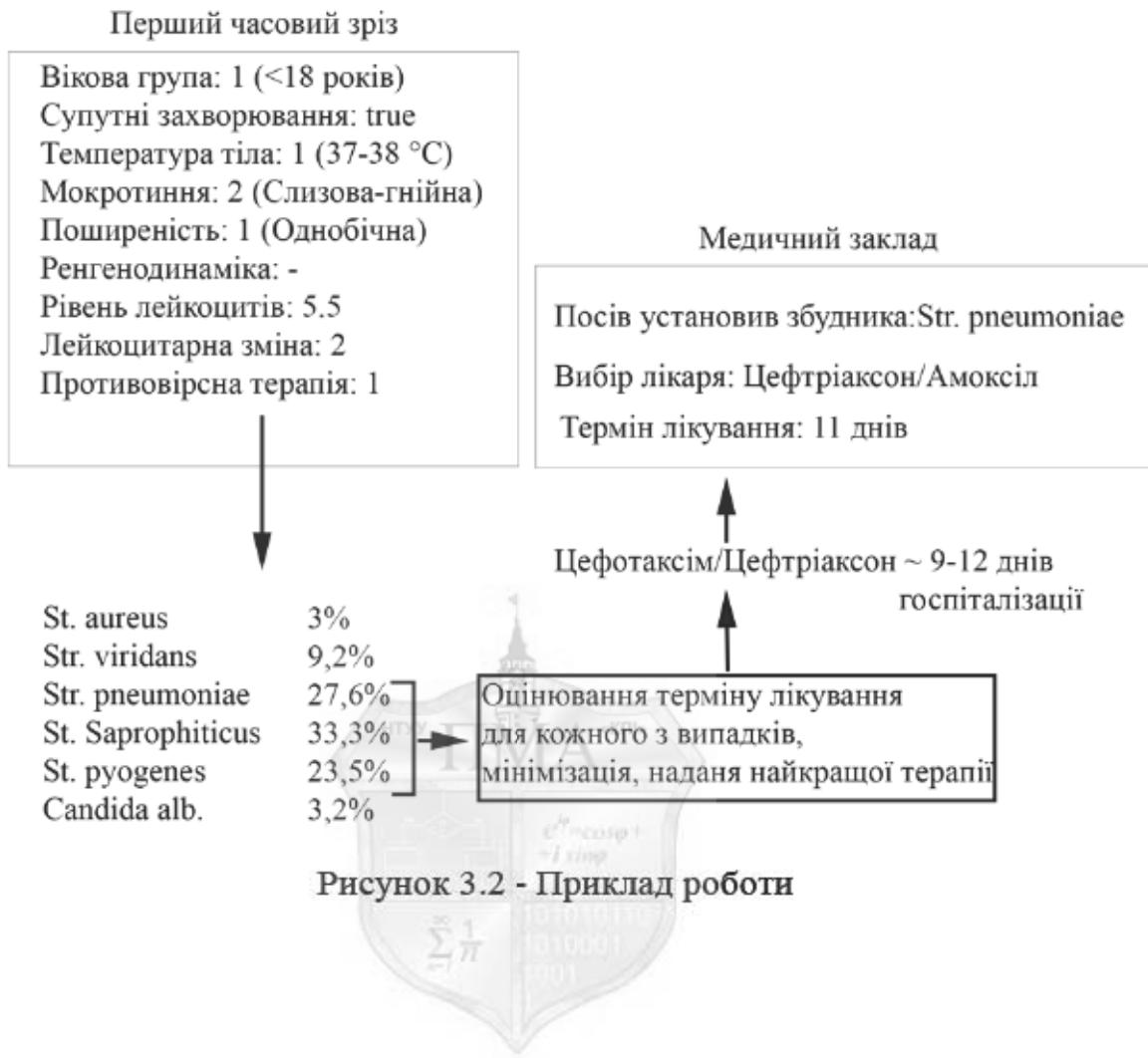
У ході аналізу предметної області було виділено наступні ключові параметри, котрі необхідно ввести для роботи системи:

- 1) Вік пацієнта
- 2) Температуру тіла
- 3) Тип мокротиння або його відсутність
- 4) Рентгенодинаміка та поширеність захворювання
- 5) Загальна кількість лейкоцитів та їх зміна
- 6) Виявлений збудник

Ці дані вводить користувач, на виході він отримує найбільш оптимальну схему терапії, в якій зазначений прогнозований термін лікування, спосіб введення та які самі лікарські засоби. Також система має надати оцінку фармацеекономічної ефективності даного лікування, зазначивши вартість зниження терміну госпіталізації, на відміну від стандартної схеми лікування.

Також було цікаво дослідити вплив додаткової противірусної терапії на термін госпіталізації. Ці дані містилися у початковій вибірці отриманої при ретроспективному аналізі історій хвороб пацієнтів, що знаходились на лікуванні в Чернігівського військового госпіталю.

Приклад виконання роботи системою зображенено на Рисунку 2.2:



### 3.1 Висновки

Було сформульовано індивідуально завдання, визначені основні пункти до розв'язання. У подальшому аналізі предметної області було отримано необхідну інформацію для подальшої роботи, список літератури та можливих консультантів у цій предметній області.

#### 4 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

Предметом аналізу було покращення наявних схем лікування хворих на поза госпітальну пневмонію, тим самим скоротивши їх термін перебування в стаціонарі, зменшивши витрати на лікування та утримання. Другорядним є фармакоекономічна оцінка отриманих схем терапії хворих на пневмонію, їх ефективність у лікуванні виражена через вартість використання по відношенню до зменшення часу госпіталізації.

Під час аналізу предметної області було проведено консультації з фахівцями лікарської справи, таким чином було виділено найбільш важливі параметри для встановлення вірної терапії, виключивши всі ті параметри, які можуть тільки обтяжити обчислення та ускладнити подальше математичне моделювання.

Такими параметрами стали наступні:

- а) Вік пацієнта
- б) Температуру тіла
- в) Тип мокротиння або його відсутність
- г) Рентгенодинаміка та поширеність захворювання
- д) Загальна кількість лейкоцитів крові та зміна лейкоцитарної формулі
- е) Виявлений збудник

Також було надано інформацію про резистентність бактерій до наявних лікарських засобів та сам перелік цих засобів. Сам перелік лікарських засобів наявний у системі має наступний вигляд:

- а) Азітроміцин
- б) Левофлоксацин
- в) Гатіфлоксацин
- г) Меропенем
- д) Цефтріаксон

- е) Амоксацилін
- ж) Пеніцилін
- з) Цефуроксім
- и) Цiproфлоксацин
- к) Сульбацеф
- л) Сульбактам
- м) Цефозалін

Кожний з даних препаратів має свою вартість та ефективність на наступні бактеріальні збудники:

- а) St. aureus
- б) Str. viridans
- в) Str. pneumoniae
- г) St. Saprofiticus
- д) St. pyogenes
- е) Candida alb.



#### 4.1. Огляд літературних джерел

За подібною тематикою було знайдено велику кількість джерел. Основними джерелами були посібники з біоінформатики та роботи за напрямком систем підтримки рішення у медичних рішеннях. Було проведено аналіз наступних літературних джерел:

##### 1) Bayesian Reasoning and Machine Learning, David Barber

Великий підручник який має на меті розібратися в Баєсових мережах для прогнозування та відповідно машинного навчання для реалізації, та роботи з неповними дата сетами.

2) Оценка тяжести и прогнозирование течения внебольничной пневмонии у лиц молодого возраста, Самойлов Р.Г.

Дисертація, в якій практично знайдено ключові фактори перебігу ВП, рекомендації щодо покращення, методи аналізу інформації.

3) Modeling physiological processes with dynamic Bayesian networks, ir. J. Hulst Робота яка розкриває процес побудови БМ, причому динамічної, котра взалежності від інформації котру вона отримує корегує параметри для кращих результатів.

4) Диагностика заболеваний методами теории вероятностей; М.Л. Жмудяк, А.Н. Повалихин, А.В. Стребуков, А.В. Гайнер, А.Л. жмудяк, Г.Г. Устинов

Праця в якій наведені приклади методів діагностики задопомоги теорії ймовірностей, наведені різні можливі варіанти рішення. Мало підходить для нашого випадку.

5) A decision support system for the diagnosis and management of pneumonia patients; Donlinik Aronsky

Дисертація в якій проводиться розробка СППР для лікування НП, відміне джерело.

#### 4.2. Огляд методів вирішення поставленої задачі

У біоінформатиці існує основні два вида вирішення задачі прогнозування невідомого:

- 1) Нейронні мережі
- 2) Баєсові мережі

Нейронні мережі - математичні моделі, а також їх програмні або апаратні реалізації, побудовані за принципом організації й функціонування

біологічних нейронних мереж – мереж первових кліток живого організму. Це поняття виникло при вивченні процесів, що протікають у мозку, і при спробі змоделювати ці процеси. Першою такою спробою були нейронні мережі Маккалока й Піттса. Згодом, після розробки алгоритмів навчання, одержувані моделі стали використовувати в практичних цілях: у завданнях прогнозування, для розпізнавання образів, у завданнях керування й ін.

ШНМ являють собою систему з'єднаних і взаємодіючих між собою простих процесорів (штучних нейронів). Такі процесори звичайно досить прості, особливо в порівнянні із процесорами, що використовуються в персональних комп'ютерах. Кожний процесор подібної мережі має справу тільки із сигналами, які він періодично отримує, і сигналами, які він періодично посилає іншим процесорам. Проте, з'єднавши їх в досить велику мережу з керованою взаємодією, такі локально прості процесори разом здатні виконувати досить складні завдання.

Нейронні мережі не програмуються у звичному змісті цього слова, вони навчаються. Можливість навчання – одне з головних переваг нейронних мереж перед традиційними алгоритмами. Технічно навчання полягає в знаходженні коефіцієнтів зв'язків між нейронами. У процесі навчання нейронна мережа здатна виявляти складні залежності між входними даними й вихідними, а також виконувати узагальнення. Це значить, що, у випадку успішного навчання, мережа зможе повернути вірний результат на підставі даних, які були відсутні в навчальній вибірці, а також неповних і/або «зашумлених», частково перекручені даних.

У контексті даного завдання НМ, абсолютно підходить, но не має можливостей точного налаштування. Переробка, додавання нових параметрів викликає труднощі. Також необхідно сказати що за досить скupoю базою вони не можуть працювати. Зокрема в нашій задачі було отримано усього 113 записів з медичних карток, у цьому випадку НМ дасть хибні значення, бо такої вибірки замало для неї.

БМ - це графічна ймовірнісна модель, що являє собою ножину змінних і їх ймовірнісних залежностей. Наприклад, баєсова мережа може бути використана для обчислення ймовірності того, що хворий пацієнт за наявністю або відсутності ряду симптомів, ґрунтуючись на даних про залежність між симптомами і хворобами.

Формально, баєсова мережа — це орієнтований ацикличний граф, вершини якого це змінні, а ребра кодують умовні залежності між змінними. Вершини можуть представляти змінні будь-яких типів, бути зваженими параметрами, прихованими змінними або гіпотезами[3]. Існують ефективні методи, які використовуються для обчислень і навчання байєсових мереж. Баєсові мережі, які моделюють послідовності змінних, називають динамічними баєсовими мережами.

#### 4.3. Вибір вирішення поставленої задачі

З наведених методів, нас будуть цікавити метод, котрий може:

- 1) Працювати з неповними даними
- 2) Дають доволі точний результат
- 3) Можуть динамічно корегувати результати, зважаючи на вхідні данні.
- 4) Дають змогу на проведення моделювань та використання тестових наборів даних

Найбільш доцільними є Баєсові мережі, тому ми використовуємо саме їх.

Недоліком нейронних мереж є їхня недетермінованість. Мається на увазі те, що після навчання є "чорний ящик", який якимось чином працює, але логіка прийняття розв'язків нейромережею зовсім схована від експерта. У принципі, існують алгоритми "витягу знань із нейронної мережі", які

формалізують навчену нейронну мережу до списку логічних правил, тим самим створюючи на основі мережі експертну систему. На жаль, ці алгоритми не вбудовуються в нейромережеві пакети, до того ж набори правил, які генеруються такими алгоритмами досить об'ємні. Тому отримання проміжних результатів, таких як один з найкращих способів терапії та її строки будуть вкрай важким завданням.

#### 4.4. Виділення параметрів з предметної області

На основі аналізу предметної області було виділено вершини, котрі відповідають певному симптуму захворювання, що має свою вагу та градацію. Сукупність вершин з симптомами та ймовірним збудником захворювання формують граф сукупного стану пацієнта зображенний на Рисунку 4.2. Виходячи з цього, система робить припущення щодо подальшої схеми терапії та приблизної тривалості госпіталізації.

Для переходу до ДБМ необхідно поєднати декілька станів хворого в певних часових інтервалах, тим самим ми можемо простежити зміну та перебіг хвороби. В моделі, що пропонується, клінічні спостереження велись на трьох проміжках часу, це надало можливість простежити зміну та перебіг захворювання в динаміці, а також проаналізувати більш детально попередні клінічні випадки.



Рисунок 4.2 - Побудована БМ для даної задачі

Формально, байєсова мережа — це спрямований ацикличний граф, вершини якого являють змінні, а ребра кодують умовні залежності між змінними. Вершини можуть представляти змінні будь-яких типів, зважені параметри, приховані змінні або гіпотезами .

На графічному відображені побудованої ДБМ еліпсами позначено спостережні симптоми у хвого на НП, схеми антибактеріальної та противірусної терапії, а також етіологічний збудник захворювання. Прямокутниками позначено: прогностичне значення тривалості госпіталізації, отриманий параметр стану хвого та значення наданої терапії. Симптоми мають градацію в балах за ступенем тяжкості, а етіологічний збудник визначається з переліку можливих. Застосовані схеми антибактеріальної та противірусної терапії мають певну ефективність проти кожного з можливих етіологічних збудників захворювання. Результатом є термін госпіталізації хвого. Наприклад стан хвого та надане лікування це детерміноване та фіксоване значення. Прогнозований час лікування- це розподіл значення кількості днів госпіталізації на осі часу.

#### 4.5 Висновки

Було встановлено основні параметри для побудови математичної моделі. Під час консультацій було вирішено проблему градації та тяжкості кожного параметру у системі.



## 5 ВИКОРИСТАННЯ БАЄСОВОЇ МЕРЕЖІ

### 5.1. Побудова баєсової мережі

На основі аналізу предметної області було виділено вершини, котрі відповідають певному симптому захворювання, що має свою вагу та градацію. Сукупність вершин з симптомами та ймовірним збудником захворювання формують граф сукупного стану пацієнта зображеного на Рисунку 5.1. Виходячи з цього, система робить припущення щодо подальшої схеми терапії та приблизної тривалості госпіталізації.

Для переходу до динамічної баєсової мережі(ДБН) необхідно поєднати декілька станів хворого в певних часових інтервалах, тим самим ми можемо простежити зміну та перебіг хвороби. В моделі, що пропонується, клінічні спостереження велись на трьох проміжках часу, це надало можливість простежити зміну та перебіг захворювання в динаміці, а також проаналізувати більш детально попередні клінічні випадки.

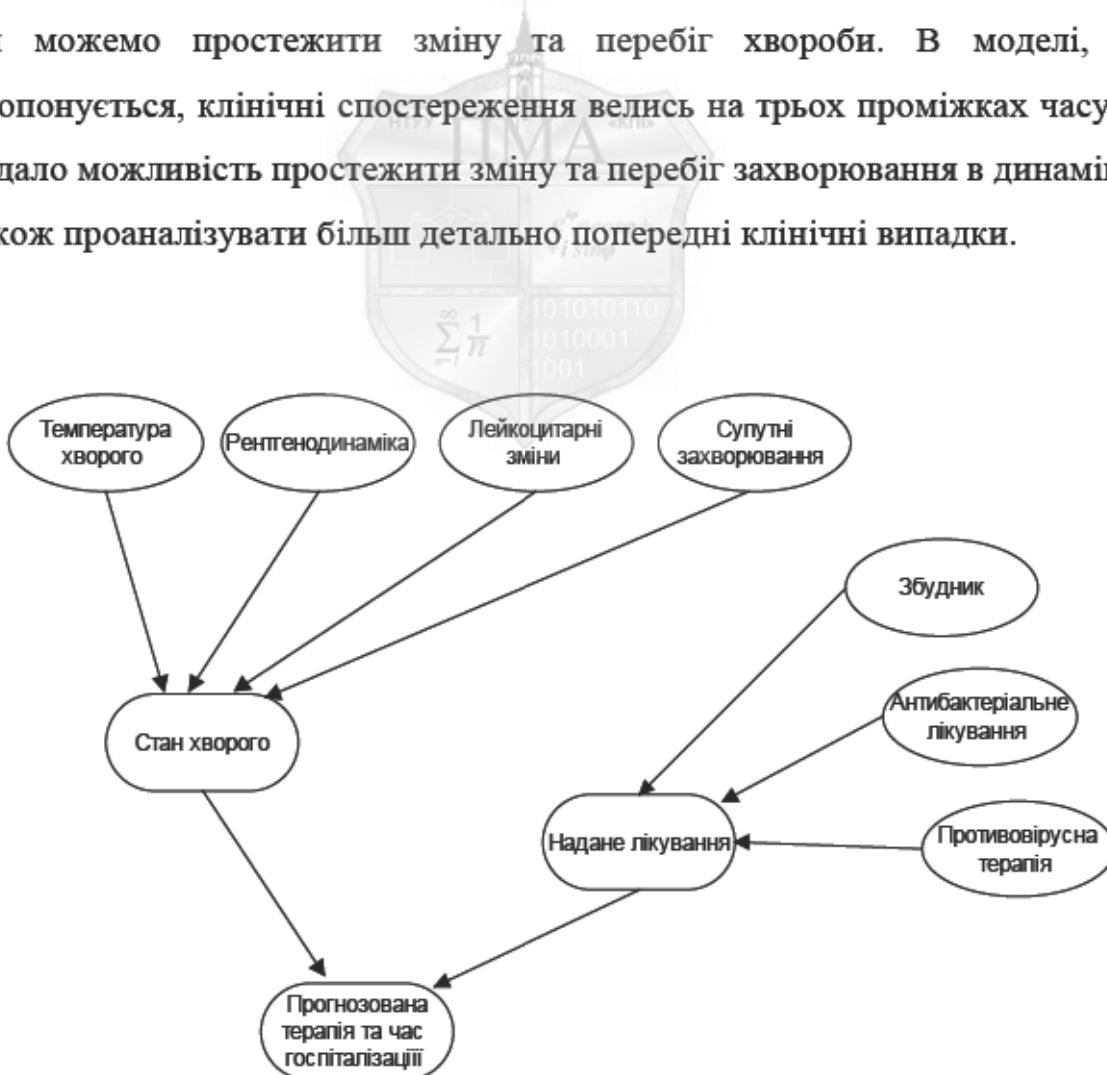


Рисунок 5.1 - Побудована БМ для даної задачі

На графічному відображені побудованої ДБМ еліпсами позначено симптоми що спостерігалися у хворого на НП, схеми антибактеріальної та противірусної терапії, а також етіологічний збудник захворювання. Прямоугольниками позначено: прогностичне значення тривалості госпіталізації, отриманий параметр стану хворого та значення наданої терапії. Симптоми мають градацію в балах за ступенем тяжкості, а етіологічний збудник визначається з переліку можливих. Застосовані схеми антибактеріальної та противірусної терапії мають певну ефективність проти кожного з можливих етіологічних збудників захворювання. Результатом є термін госпіталізації хворого представлених у ліжко-днях.

## 5.2. Опис алгоритму обчислення ймовірностей у динамічних баєсових мережах



Згідно літературних джерел обрахунок ймовірностей всередині БМ проводиться наступними кроками [4]:

- Аналіз існуючих подібних випадків на даному часовому проміжку.
- Розрахунок ймовірності того чи іншого збудника, що є невідомим на початок лікування, та ймовірності симптомів на основі ретроспективних даних про хворого.
- Аналіз усіх можливих варіантів та розрахунок ймовірності кожного значення параметру на основі теоремі Баєса, яка має наступний вигляд:

$$p(X|Y) = \frac{p(Y|X)*p(X)}{p(y)}, \text{ в даному вигляді вона не може бути застосована,}$$

тому існує наступний вигляд:  $p(X|d) = \sum_i \alpha \cdot p(X|h_i) \cdot p(d|h_i) \cdot p(h_i)$ ,

де  $p(X|d)$  шукана вірогідність при виконані умови  $d$ ,  $X$  фіксований випадок з переліку можливих, за умов відомих значень вірогідностей при усіх можливих гіпотезах  $h_i$ . Розрахунок виходячи з існуючих даних  $p(h_i)$ ,  $p(d|h_i)$ ,  $p(X|h_i)$ . Розрахунок нормуючого коефіцієнту  $\alpha$ .

- г) Якщо складність завелика, необхідно зробити декомпозицію БМ, для більш легкого обрахунку.
- д) Перехід до ДБМ, з використанням нових значень стану пацієнта.
- е) Отримання результату із ЕМ-алгоритму .
- ж) Розрахунок орієнтованого терміну лікування, та різниці між прийомом противірусних препаратів та без них.

Існують два види навчання: параметричне та структурне. У даній роботі було обрано параметричне навчання, яке ґрунтується на аналізі вхідних даних. У даному випадку це аналіз усіх подібних медичних карт з анамнезами, цей підхід дає можливість працювати з неповними даними. Одним із можливих рішень параметричного навчання є ЕМ-алгоритм, який і був використаний.

Тому на першому етапі ми знаходимо ймовірність, того чи іншого параметру у вершині, далі будуємо функцію максимізації спостереження, для цього необхідно побудувати наступну функцію[5]:

$$\langle l_c(\theta; x, z) \rangle_q \triangleq \sum_z q(z|x, \theta) \cdot \log p(x, z|\theta) \quad (5.1)$$

Де,  $p(x, z|\theta)$  ймовірність відомого  $x$  та невідомого  $z$  випадку, за параметрів  $\theta$ , де  $q(z|x)$  ймовірнісна залежність між відомою, та невідомою величиною. Наступним етапом є знаходження:

$$\begin{aligned} l(\theta; x) &= \log p(x|\theta) = \log \sum_z q(z|x) \cdot \frac{p(x, z|\theta)}{q(z|x)} \\ &\geq \sum_z q(z|x) \cdot \log \frac{p(x, z|\theta)}{q(z|x)} \triangleq L(q, \theta, x) \end{aligned} \quad (5.2)$$

І далі проводиться максимізація цієї функції по одному параметру, наступним чином (EM-алгоритм):

$$q^{t+1} = \arg \max_q L(q, \theta^t, x) \quad (5.3)$$

$$\theta^{t+1} = \arg \max_{\theta} L(q^{t+1}, \theta, x) \quad (5.4)$$

На виході ми найбільш очікуване значення, невідомого параметру, в нашому випадку це кількість днів госпіталізації.

Існують інші способи навчання такі як структурне, але воно повністю спирається на дані у базі та нагадує собою неглибокі нейроні мережі.

### 5.3 Висновок

Було проаналізовано метод на основі побудови басової мережі. Знайдені необхідні робочі формули та етапи розв'язання всередині математичної моделі. Обрано спосіб навчання від вибірки, а саме параметричний спосіб, що дозволяє використовувати невідомі параметри.

## 6 ФАРМАКОЕКОНОМІЧНА ЕФЕКТИВНІСТЬ ТЕРАПІЇ

### 6.1. Фармакоекономічна ефективність

Фармакоекономіка - нова самостійна наука, яка вивчає в порівняльному плані співвідношення між витратами та ефективністю, безпекою застосування препарату, якістю життя при альтернативних схемах лікування ( профілактики ) захворювання.

Комплексний підхід до оцінки доцільності застосування медичних технологій припускає взаємопов'язану оцінку наслідків (результатів) і вартості медичних втручань. Найбільш важливим у цьому визначенні є саме взаємопов'язана оцінка, тобто мова йде не просто про порівняння витрат, а про оцінку співвідношення між витратами і отриманими результатами.

З точки зору практичного лікаря це означає, що фармакоекономіка - це не пошук найбільш дешевих лікарських засобів і виправдання їх використання, а розрахунок витрат, необхідних для досягнення бажаної ефективності, і співвіднесення цих витрат з можливостями.

### 6.2. Фармакоекономічна модель у даній задачі

В даній задачі фармакоекономічна модель дозволяє обирати ту терапію котра надає найкраще співвідношення ефективності та ціни, при умові рівності з іншими. Також ця модель дозволяє оцінити ефективність кожного з можливих препаратів для лікування НП [6].

У ході аналізу предметної області було побудовано наступну модель, переходи познають зміну стану, та вибір препарату/терапії, дана модель зображена Рисунку 6.1:

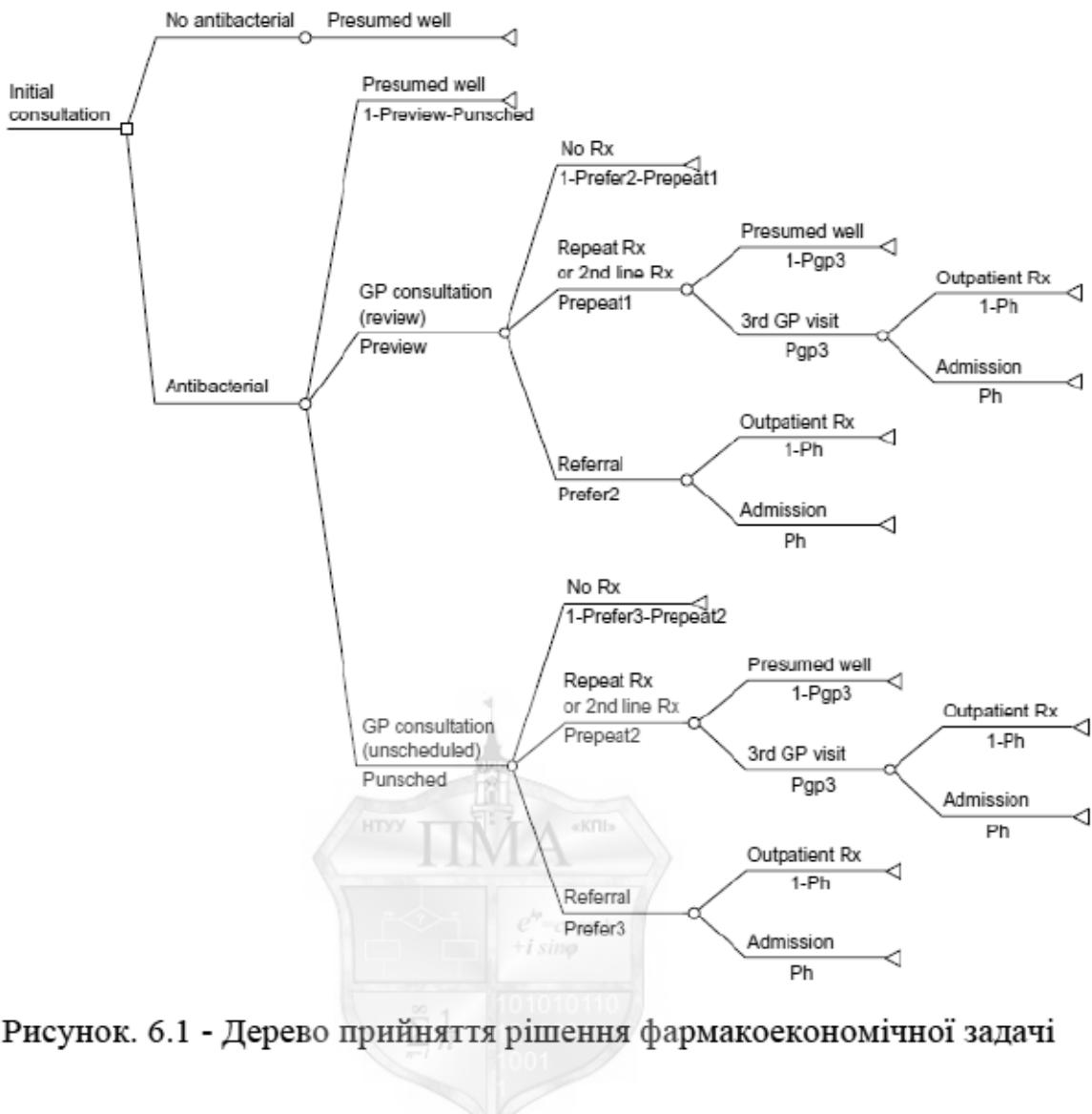


Рисунок. 6.1 - Дерево прийняття рішення фармакоекономічної задачі

Для отримання вартості необхідно в кінцях (на листях) вказати ціну даної терапії, порахувавши вартість ліків наданих при переходах по дереву.

А результат терміну госпіталізації для кожного з видів прийому має бути вирахуваний у БМ, враховуючи кожну особливу терапію.

Глибину дерева можна задати кількістю каскадів у прийомі ліків. А ширина виходить з можливих наявних ліків.

### 6.3 Висновок

Було проведено аналіз предметної області у сфері фармакоекономіки та побудовано математичну модель для оцінки ефективності наданої терапії, котра визначається як перевага заданої терапії у кількості днів госпіталізації перед номінальною терапією та вартістю між ними.



## 7 ОБРАННЯ ШЛЯХУ ПРОГРАМНОЇ РЕАЛІЗАЦІЇ

У ході аналізу поставленої задачі було отримано наступні вимоги:

- а) Програмне реалізація має виконувати вхід та вихід користувача зі системи.
- б) Час обчислень має складати невеликий проміжок часу, та мало залежити від можливостей користувача.
- в) Бути кроссплатформеним рішенням.
- г) Дозволяє зняти навантаження обчислень з користувача.
- д) Дозволяє розробити ‘тонкий’ кліент, тобто такий що лише відображає отриману інформацію та надсилає необхідну інформацію на сервер.

Наступним кроком було обрання засобів для реалізації. В силу можливостей, було обрано розробити проект на основі зв’язки наступних технологій:

- а) Python 2.7.6, інтерпретуюча мова програмування, на котрій буде запрограмовано логіку та керування програмним продуктом.
- б) MySQL, СУБД та БД для керування та збереження всіх необхідних даних.
- в) Werkzeug, засіб для створення локального сервера, для подальшої розробки на основі ПК розробника.
- г) HTML та CSS, мова розмітки веб-сторінок та її таблиця стилів, для створення інтерфейсу користувача.
- д) Jinja 2, засіб для динамічного обновлення вмісту HTML та керування вигляду.
- е) Flask API, засіб для роутингу на сервері між сторінками та передачі інформації між застосунком Python та веб-сторінки.
- ж) Navicat for MySQL, засіб для зручного адміністрування змісту БД MySQL.

## 7.1 Висновки

Було обрано інструменти для побудови заданої системи, обрано кожну з необхідних ланок для побудови цілісного проекту. Проаналізовано переваги кожної технології для їхнього використання.



## 8 ПРОГРАМНА РЕАЛІЗАЦІЯ

Було створено з перелічених засобів клієнт-серверний застосунок.

Це дозволить досягнути розділення нашого застосунки на три рівня, котрий з який може змінюватися та доповнятися, мало впливаючи на інші, тим самим зменшивши час на обслуговування . Наступні три рівні:

- a) Рівень представлення даних, який по суті являє собою інтерфейс користувача і відповідає за представлення даних користувачеві і введення від нього керуючих команд, буде представлено у вигляді Веб-інтерфейсу, подібне рішення дозволяє досягти кроссплатформеності та дозволить використовувати її на будь-яких пристроях котрі мають веб-браузер. Основні технології котрі ми застосуємо: HTML, CSS, JScript.
- b) Прикладний рівень, який реалізує основну логіку застосунку і на якому здійснюється необхідна обробка інформації, котра виконується на сервері тим самим дозволить не зважати на обчислювальні можливості у клієнта. Ця частина була виконана на Python.
- v) Рівень управління даними, який забезпечує зберігання даних та доступ до них. Виконаний на основі MySQL. Архітектура даних котрого має наступний вигляд зображений на Рисунку 8.1:

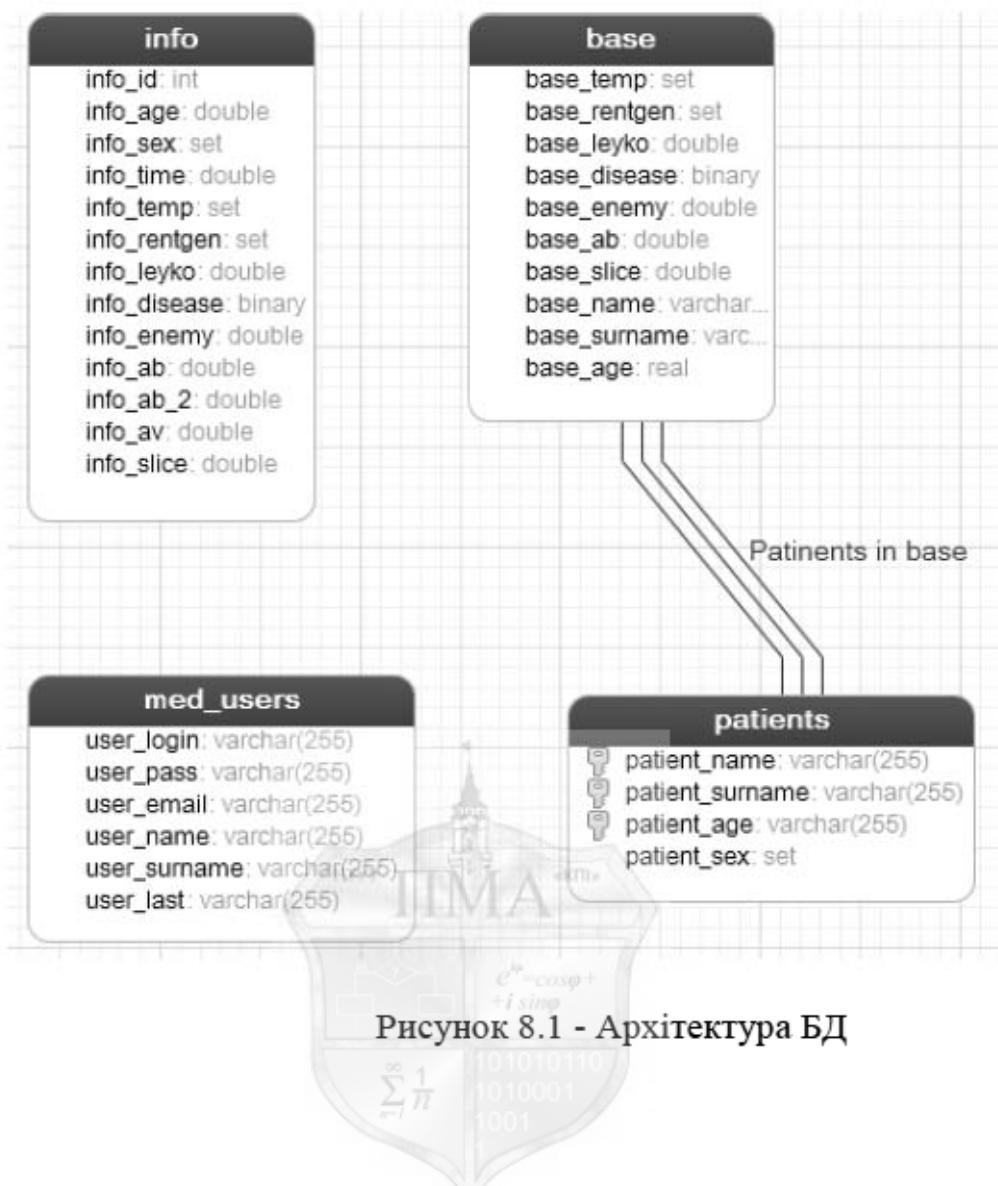


Рисунок 8.1 - Архітектура БД

У таблиці *info* зберігаються усі наявні випадки перебігу НП. Таблиця *med\_user* відповідає за зарегестрованих користувачей. Таблиця *patients* містить лише уведені нові випадки захворювання. А таблиця *base* відповідає за збереження даних про нові випадки, і по цим даним проводиться аналіз.

Було створено наступну ієрархії веб-сторінок на сервері зображену на Рисунку 8.2:

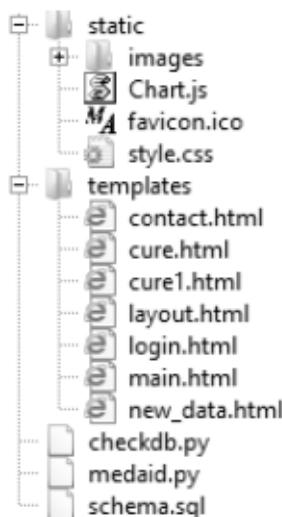


Рисунок 8.2 - Ієрархія веб-засобу

- a) Images – містить усі графічні елементи сторінок.
- б) Chart.js – виконує функцію зображення необхідних даних у вигляді графіку.
- в) Папка Templates та файли в них мають в собі необхідні сторінки веб-застосунки сторінок.
- г) Medaid.py – сам Python застосунок, який відповідає за логіку переходів між сторінками та обчислення.

Повний код програми та пояснення її елементів наведені в Додатках.

## 8.1 Робота у програмі

Для входу необхідно набрати у веб-браузері відповідну адресу, у тестовому режимі для цього було необхідно набрати: <http://localhost:5000>,

після цього необхідно зайти до системи використовуючи надані дані та ввести у форму зображену на Рисунку 8.3:



Рисунок 8.3 – Вхід до сервісу

Після цього буде відкрито наступну сторінку, на якій необхідно ввести дані про пацієнта зображені на Рисунку 8.4, такі як:

- а) Ім’я хворого
- б) Прізвище хворого
- в) Стать хворого
- г) Рік народження хворого
- д) Температура хворого
- е) Рентгенодинаміка хворого
- ж) Лейкоцитарні зміни хворого
- з) Наявність супутних захворювань
- и) Збудник за його наявності
- к) Та час коли буди введені ці дані

The screenshot shows the MEDAID software interface. At the top, there is a navigation bar with tabs: Головна (Home), Пациєнти (Patients), Лікування (Treatment), and Робота з даними (Data Work). The main area is titled 'ВВЕДІТЬ ДАНІ ДЛЯ НОВОГО ПАЦІЕНТА' (Enter data for a new patient). The form contains the following fields:

- ІМ'Я (Name): Іван квітка
- Призвище (Surname): Прізвище хворого
- Стать (Gender): Чоловік
- Рік народження (Year of birth): 1980
- Температура (Temperature): 37.5
- Рентгенодинаміка (RADIOLOGY): Повністю збережена
- Лейкоцитарні зміни (Leukocyte changes): Кількість лейкоцитів-ліпінів
- Супутні захворювання (Associated conditions):  Не має  Належні
- Збудник (Pathogen): Не виявлено
- Проміжок часу (Time interval): Тільки прибув

Рисунок 8.4 – Ввід даних

Далі буде переадресовано на наступну сторінку де буде зображеного результат, а саме прогнозований час та препарат для оптимальної терапії. Графік зображений на Рисунку 8.5, представляє розподілення часу госпіталізації для терапії :

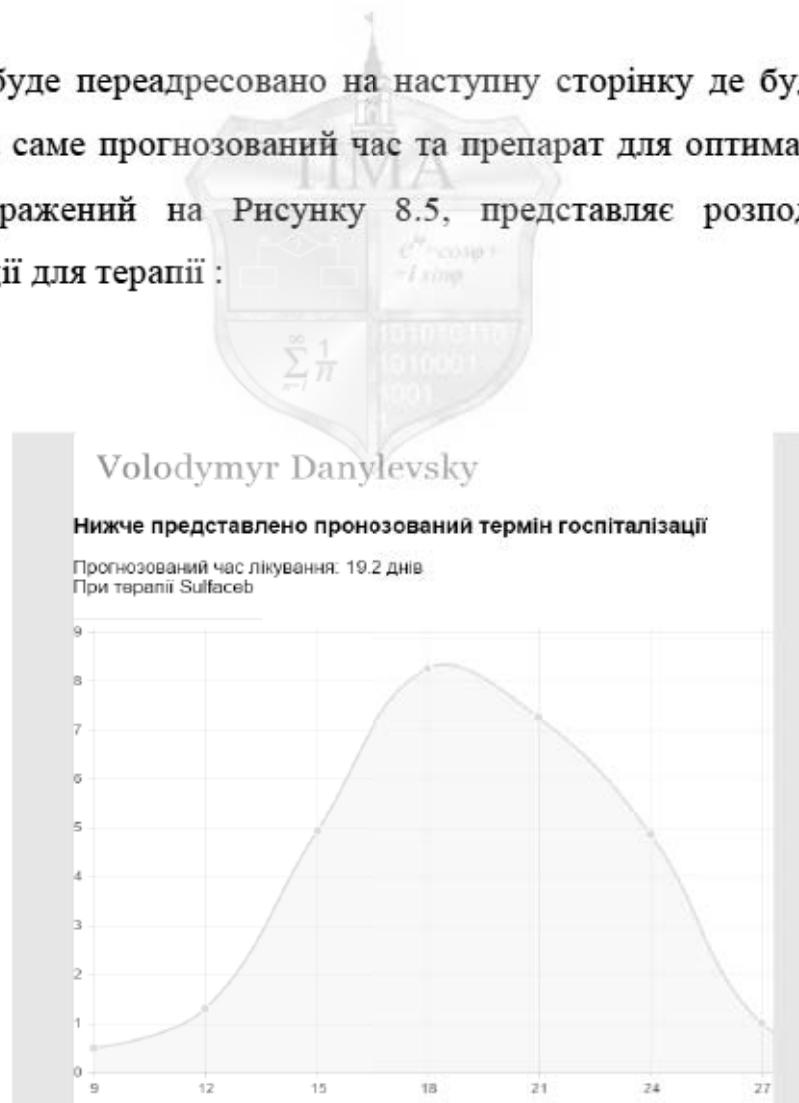


Рисунок 8.5 – Результат обчислення

## 8.2 Висновки

Було збудовано інтерфейс користувача, який відповідає заявленим вимогам. Він має всі необхідні перевірки та дозволяє легко ввести нові дані для користувача до системи.



## 9 ПРОГНОЗУВАННЯ

Було проведено аналіз існуючих випадків отриманих з Чернігівського військового госпіталю та результатів сгенерованих програмою, загальна вибірка з 114 випадків середній прогнозований термін госпіталізації склав 19,28 доби, при практичному терміні госпіталізації 21,3 доби.

Це обумовлено іншим вибором програми лікарського засоба, який може теоретично зменшити час госпіталізації.

Також було досліджено вибірку людей котрі пройшли противірусну терапію у порівняні з тими що не пройшли. Середній очікуваний термін госпіталізації для тих хто проходив противірусну терапію склав 15,85 доби, за практичного результата в 18,9 доби.



## ВИСНОВКИ

У роботі розглянуті основні способи прогнозування у біонформатиці, та детально розібрано спосіб прогнозування на основі побудови баєсової мережі. Було проведено аналіз предметної області та виділені необхідні параметри баєсової мережі. Виходячи зі специфіки кожного параметру було обрано параметричне навчання зі всіх можливих.

Було збудовано необхідну модель баєсової мережі, проведено спрощення існуючої моделі завдяки d-сепарації, що знизило обчислювальні складності

Було збудована система для прогнозування часу госпіталізації та знаходження оптимальної терапії. До неї було внесені необхідна вибірка для навчання та проведено параметричне навчання.

Проведено аналіз практичних та теоретичних отриманих зі системи даних. У порівняні з даними отриманими на випадків негоспітальної пневмонії та результатами системи отримана похибка для часу госпіталізації склала 10,7%, відповідно середній термін госпіталізації хворих з отриманих даних Чернігівського військового госпіталю склав – 21,3 день, в той час система надала відповідь у 19,23 дня.

У подальшому подібна система має можливість для подальшої інтеграції у системи медичних закладів, тим самим забезпечивши лікарський склад інструментом для аналізу ефективності ліків та можливістю прогнозувати час госпіталізації пацієнтів.

Основні положення дипломної роботи викладені у вигляді тез на наукові конференції «Прикладна математика та комп’ютинг 2015».

## СПИСОК ЛІТЕРАТУРИ

1. Barber D. Bayesian Reasoning and Machine Learning / Barber D. - DRAFT, 2010 – 610 с.
2. Мусина В.Ф. Байесовские сети доверия как вероятностная графическая модель для оценки медицинских рисков / Мусина В.Ф. - СПИРАН, 2013 – 17 с.
3. Hulst J. Modeling physiological processes with dynamic Bayesian networks / Hulst J. – Delft University of Techology, 2006 – 180 с.
4. Aronsky D. A decision support system for the diagnosis and management of pneumonia patients / Aronsky D. – University of Utah, 2001 – 112 с.
5. Bonten M. Bayesian network models for the management of ventilator-associated pneumonia / Bonten M. – University of Utrecht, 2008 – 185 с.
6. Бідюк П.І., Демківський Є.О. Аналіз даних з використанням байєсівських моделей / Бідюк П.І., Демківський Є.О. - Наукові вісті НТУУ "КПІ" 2012 – 15 с.