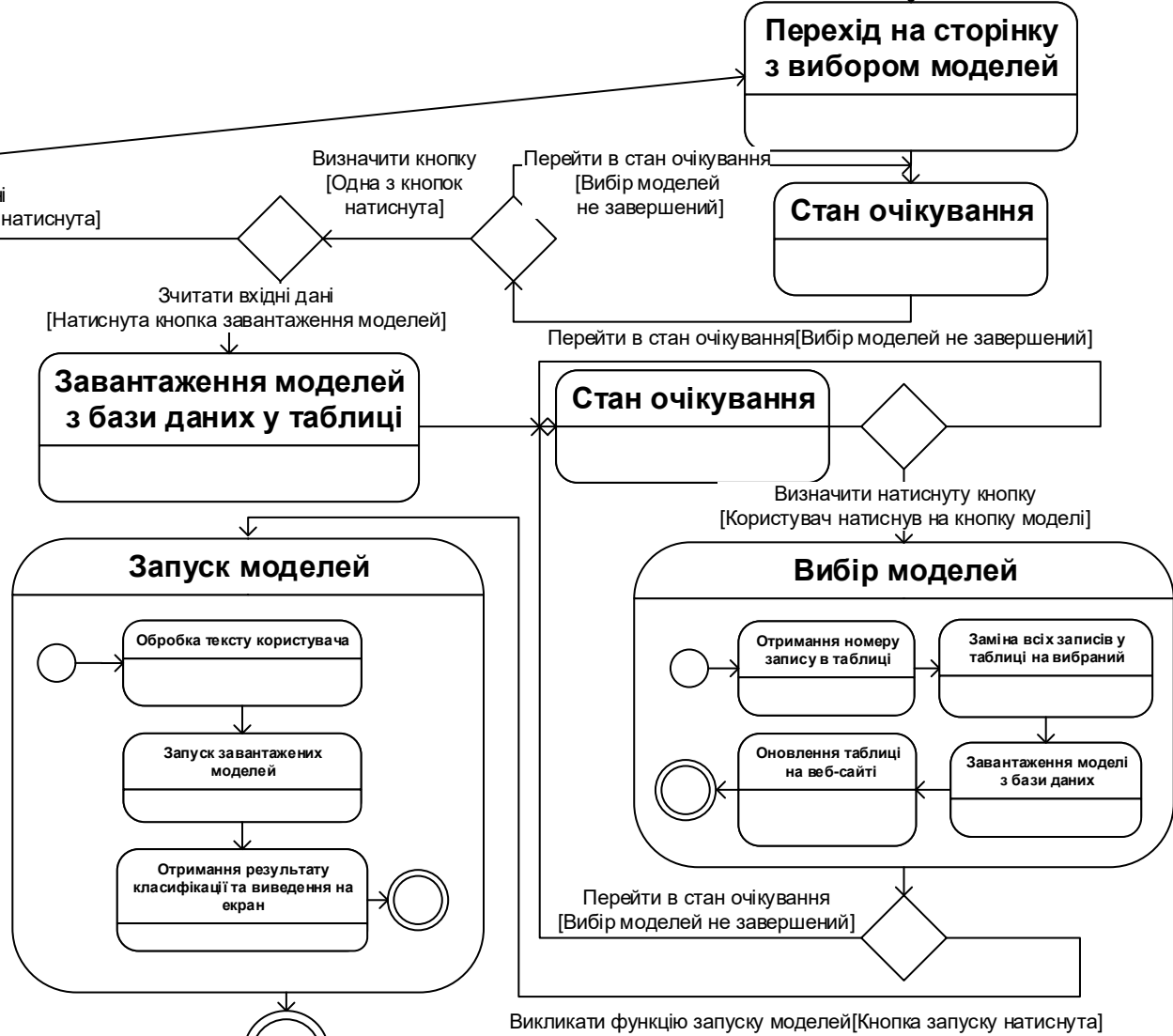
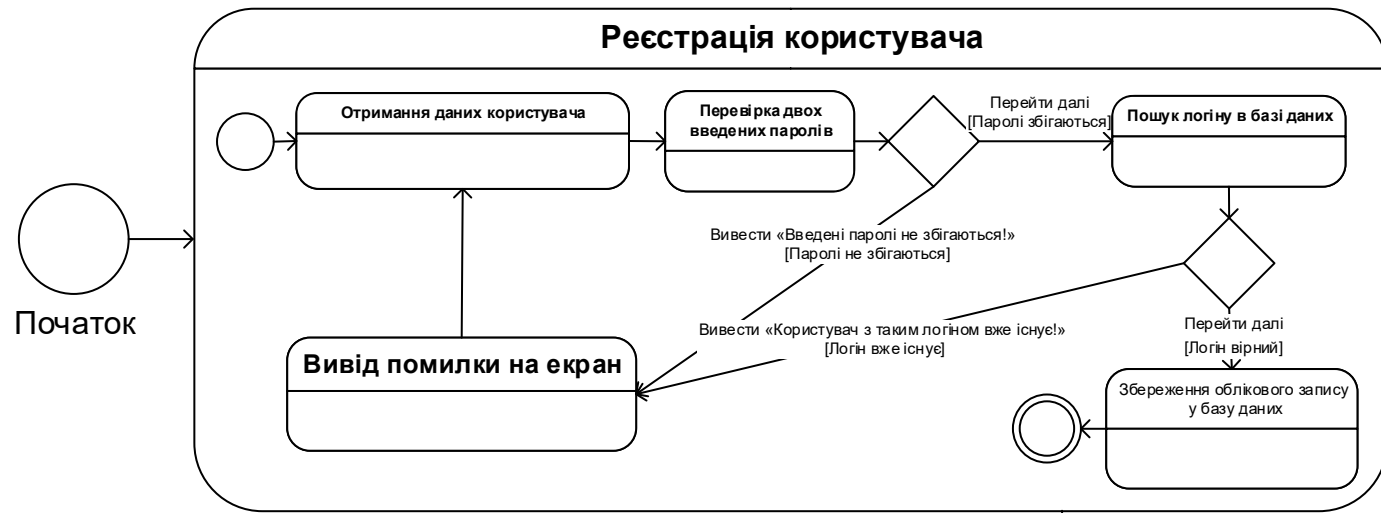
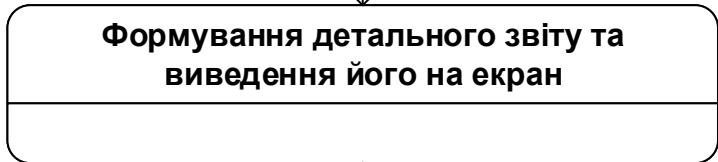
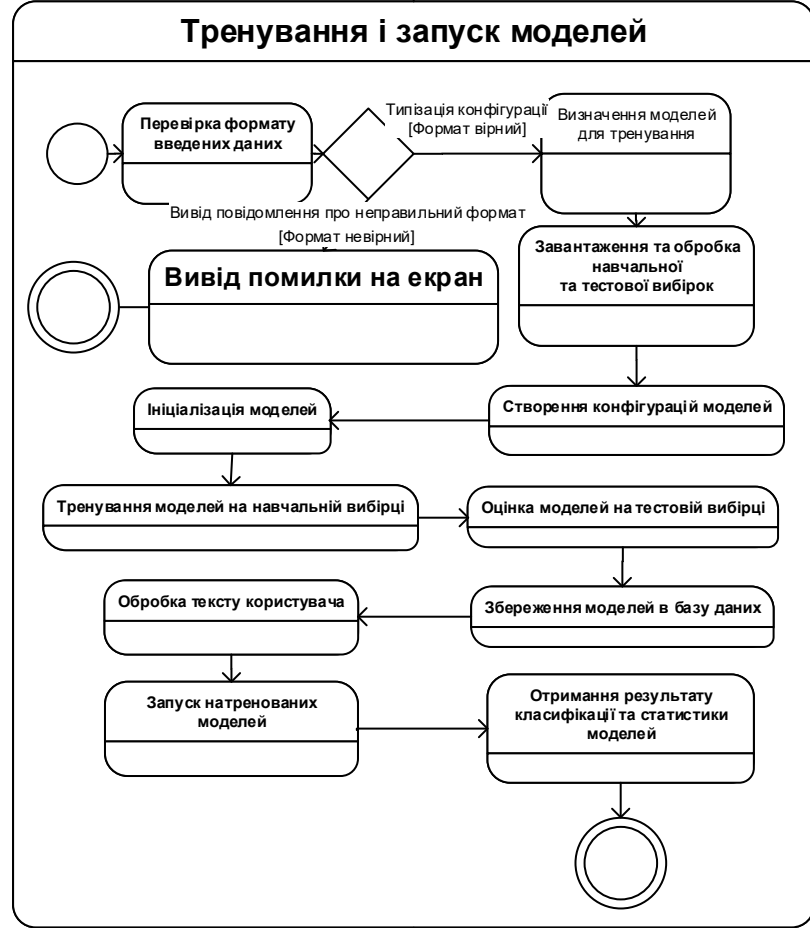
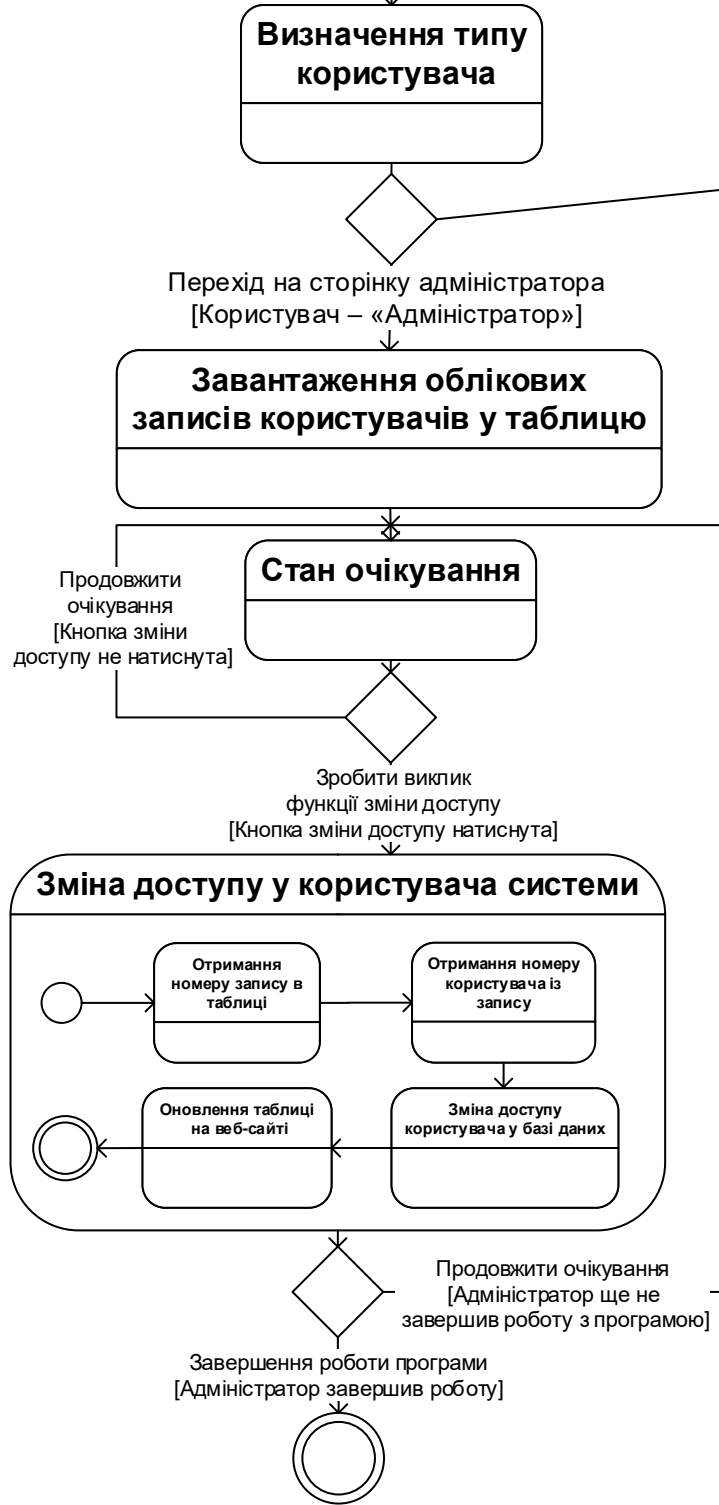


					ДП 6303.02.000 ССВ					
					Схема структурна варіантів використання					
Зм.	Арк.	№ документа	Підпис	Дата	Інформаційна технологія визначення тональності текстів					
Розробив		Блінков Є.М.								
Перевірив		Жданова О.Г.								
Т. кон.										
					КПІ ім. Ігоря Сікорського кафедра АСОІУ гр. ІС-63					
Н. кон.		Проскура С.Л.								
Затвердив		Жданова О.Г.								

Літера			Маса		Масштаб	
Аркуш 1			Аркушів 1			



						ДП 6303.03.000 ССС					
						Схема структурна станів		Літера		Маса	Масштаб
Зм.	Арк.	№ документа	Підпис	Дата		Інформаційна технологія визначення тональності текстів		Аркуш 1		Аркушів 1	
Розробив	Блінков Є.М.										
Перевірив	Жданова О.Г.										
Т. кон.											
Н. кон.	Проскура С.Л.										
Затвердив	Жданова О.Г.										

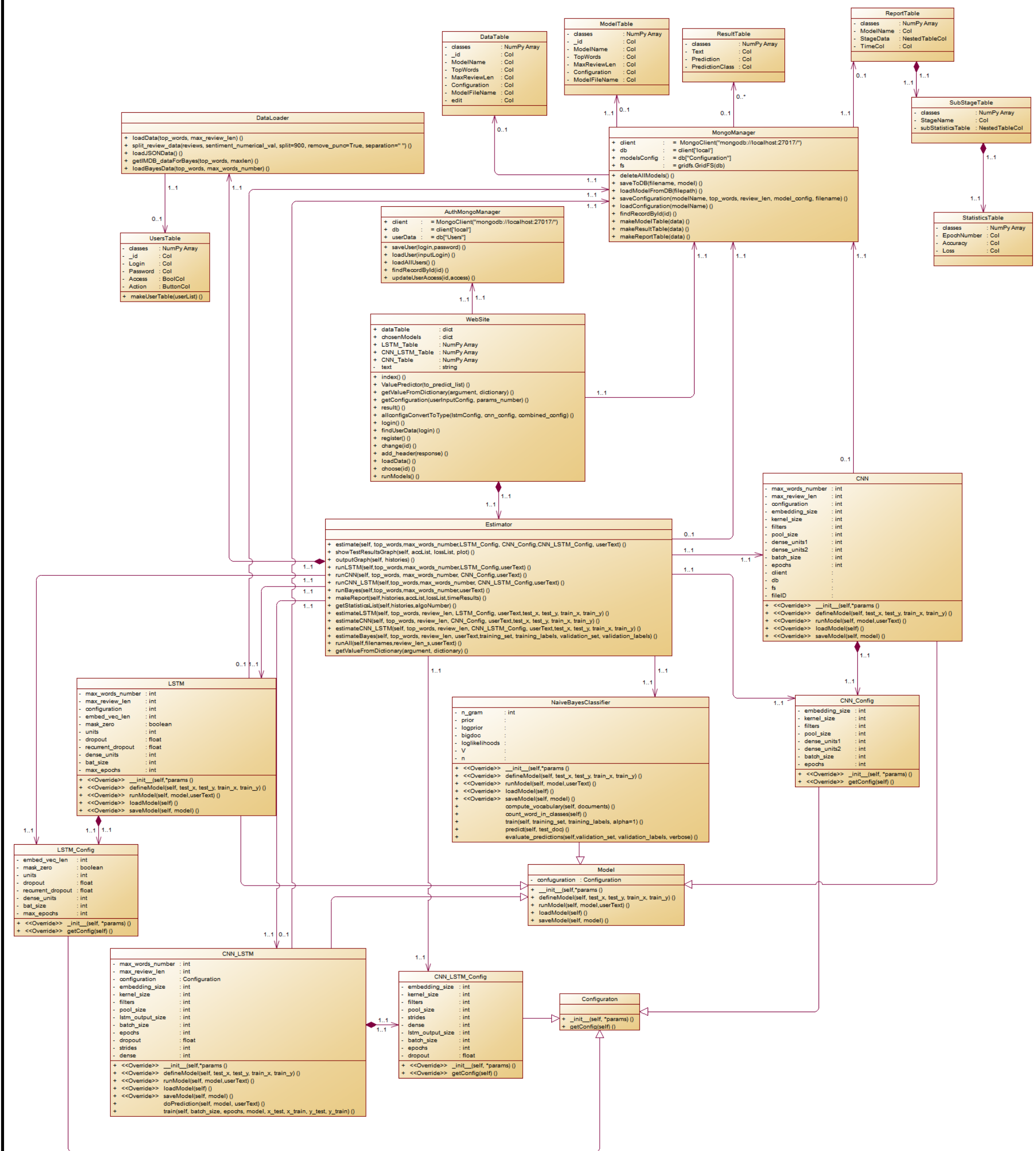
Configuration	
<u>_id</u>	<u>ObjectID</u> <pk>
ModelName	String
TopWords	Int32
MaxReviewLength	Int64
Configuration	Array
ModelFileName	ObjectID <fk>

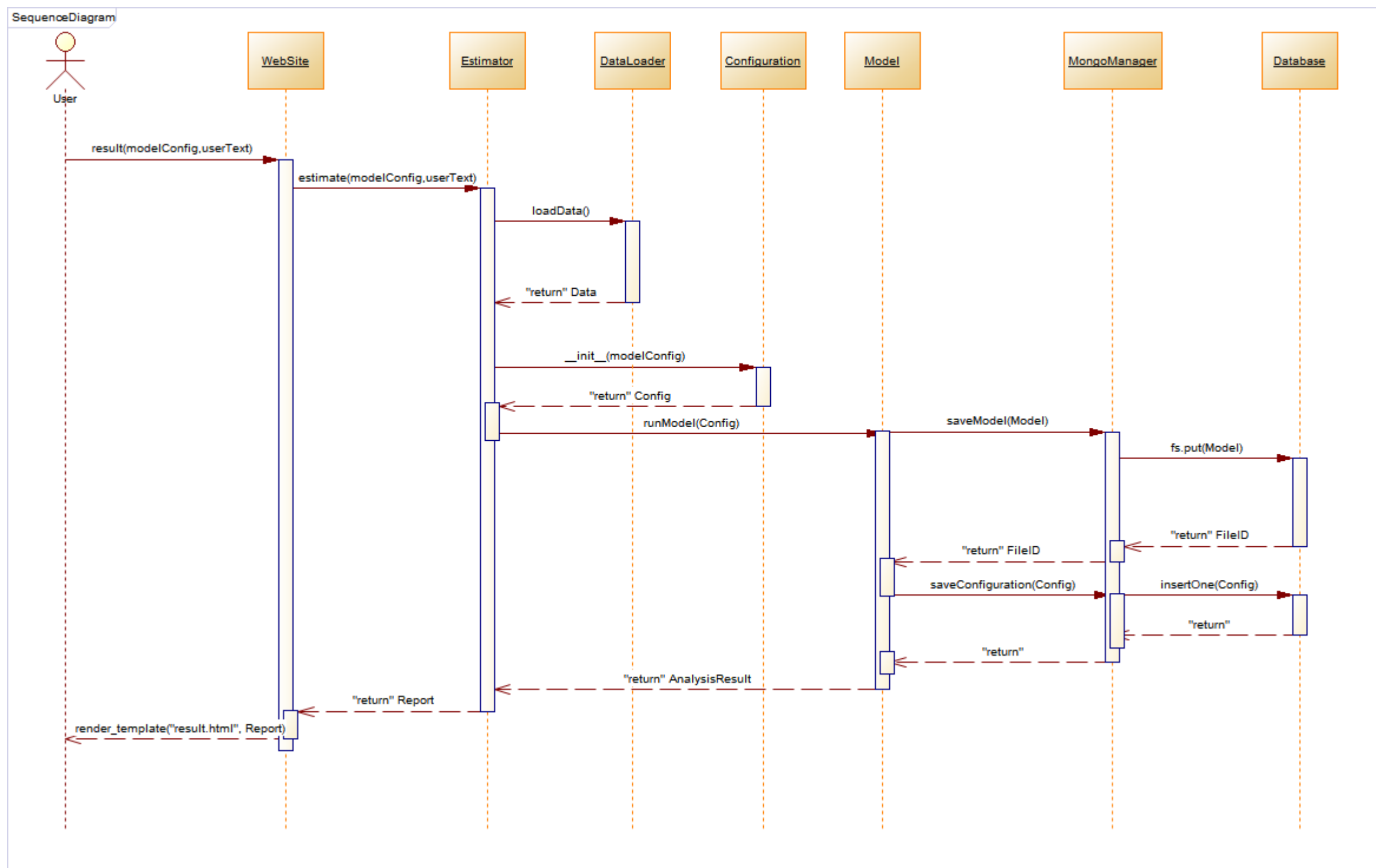
FK_CONFIGUR_REFERENCE_FS.FILES

fs.files	
filename	String
length	Int64
Content_type	String
Upload_date	DATE
<u>md5</u>	<u>ObjectID</u> <pk>
metadata	String

Users	
<u>_id</u>	<u>ObjectID</u> <pk>
Login	String
Password	String
Access	Bool

					ДП 6303.04.000 СБД				
					Схема бази даних		Літера	Маса	Масштаб
Зм.	Арк.	№ документа	Підпис	Дата					
Розробив		Блінков Є.М.							
Перевірів		Жданова О.Г.							
Т. кон.							Аркуш 1		Аркушів 1
					Інформаційна технологія визначення тональності текстів		КПІ ім. Ігоря Сікорського кафедра АСОІУ гр. ІС-63		
Н. кон.		Проскура С.Л.							
Затвердив		Жданова О.Г.							





					ДП 6303.06.000 ССП			
					Схема структурна послідовності	Літера	Маса	Масштаб
Зм.	Арк.	№ документа	Підпис	Дата				
Розробив		Блінков Є.М.						
Перевірів		Жданова О.Г.						
Т. кон.					Інформаційна технологія визначення тональності текстів	Аркуш 1		Аркушів 1
Н. кон.		Проскура С.Л.				КПІ ім. Ігоря Сікорського кафедра АСОІУ гр. ІС-63		
Затвердив		Жданова О.Г.						

Рішення з математичного забезпечення

Схема алгоритму роботи рекурентної нейронної мережі з технологією довгої короткочасної пам'яті (ДКЧП)

Крок 1: За допомогою Word2Vec з кожного слова тексту

утворити векторне представлення

Крок 2: Отримані слова $\{T_1, \dots, T_n\}$ один за одним ввести у

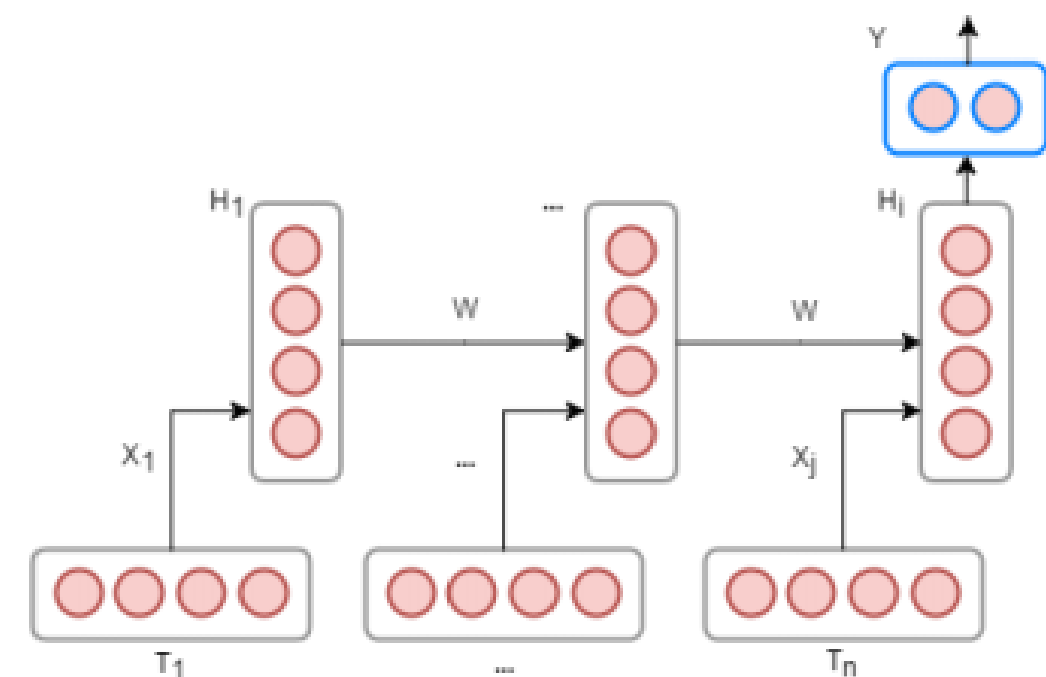
ДКЧП

Крок 3: В кожен момент часу j вихід W прихованого шару

передати назад у прихований шар H_i разом з наступним вхідним вектором x_{j+1} у наступний момент часу $j+1$.

Крок 4: Останній вихід W передати вихідному шару

нейронної мережі



Демонстраційний плакат до дипломного проекту

„Інформаційна технологія визначення тональності текстів”

Виконав студент гр. ІС-63

Блінков Є.М..

Керівник ДП

Жданова О.Г.

Веб-сторінка з вибором моделей для їх навчання

Виберіть модель визначення тональності тексту для її навчання

Список моделей:	Конфігурація моделі:	Опис необхідних параметрів:
<input checked="" type="checkbox"/> Наївний Баєсів класифікатор		
<input type="checkbox"/> Нейронна мережа з ДКЧП	32,True;100;0.2;0.2;1;1	1.Довжина векторного представлення тексту, 2.Маска для нульових елементів векторного представлення тексту, 3.Кількість нейронів у шарі ДКЧП,
<input type="checkbox"/> Комбінована нейронна мережа(ДКЧП+ЗНМ)	32;3;32;2;100;32;5;0.01	1.Довжина вихідних векторів з шару Embedding, 2.Розмір вікна згортки, 3.Кількість фільтрів для аналізу векторів слів,
<input type="checkbox"/> Згорткова нейронна мережа (ЗНМ)	32;3;32;2;250;1;128;3	6.Кількість нейронів у другому звичайному шарі нейронної мережі, 7.Розмір вибірки, для здійснення однієї ітерації тренування моделі, 8.Кількість періодів навчання нейронної мережі

Кількість найпоширеніших слів у словнику:

5000

Кількість слів у текстах:

100

Введіть текст для визначення його тональності:

The movie was awesome. I love it

Почати навчання

Завантажити дані з бази даних

Веб-сторінка з вибором моделей завантажених з бази даних

Виберіть одну з моделей для проведення аналізу текстів на їх тональність

Моделі Рекурентної нейронної мережі з ДКЧП :

Назва моделі	Кількість найчастіших слів	Довжина тексту	Конфігурація моделі	Назва файлу з моделлю	Вибрати модель
Istm_model	5000	100	[32, True, 100, 0.2, 0.2, 1, 128, 3]	5ec90a460dd9443be84e9fe5	<button>Вибрати модель</button>
Istm_model	5000	100	[32, True, 100, 0.2, 0.2, 1, 128, 3]	5eca83040732a5322f80c7b9	<button>Вибрати модель</button>
Istm_model	7000	100	[32, True, 100, 0.2, 0.2, 1, 128, 3]	5eca89f059749cb5748a48cb	<button>Вибрати модель</button>
Istm_model	5000	100	[32, True, 100, 0.2, 0.2, 1, 128, 3]	5ecabeb46a1beadc2431478f	<button>Вибрати модель</button>
Istm_model	5000	100	[32, True, 100, 0.2, 0.2, 1, 128, 3]	5ecacadc8e380e6adedeb3a	<button>Вибрати модель</button>
Istm_model	5000	100	[32, True, 100, 0.2, 0.2, 1, 128, 3]	5ecac9ec76f54de69f7469580	<button>Вибрати модель</button>

Моделі Згорткової нейронної мережі :

cnn_model	5000	100	[32, 3, 32, 2, 250, 1, 128, 3]	5ecac50462fbeb665bddbe71	Вибрати модель
cnn_model	5000	100	[32, 3, 32, 2, 250, 1, 128, 3]	5ecac6152473d000afe899a7	Вибрати модель
cnn_model	5000	100	[32, 3, 32, 2, 250, 1, 128, 3]	5ecac7c7038f0a34d089ec5a	Вибрати модель
cnn_model	5000	100	[32, 3, 32, 2, 250, 1, 128, 3]	5ecac883fd20c4282947721f	Вибрати модель
cnn_model	5000	100	[32, 3, 32, 2, 250, 1, 128, 3]	5ecaca39ae8380e6adedeb1d	Вибрати модель
cnn_model	5000	100	[32, 3, 32, 2, 250, 1, 128, 3]	5ecace2e6f54de69f7469563	Вибрати модель

Моделі Комбінованої нейронної мережі:

Вигляд веб-сторінки з вибраними моделями

Виберіть одну з моделей для проведення аналізу текстів на їх тональність

Моделі Рекурентної нейронної мережі з ДКЧП :

Назва моделі	Кількість найчастіших слів	Довжина тексту	Конфігурація моделі	Назва файлу з моделлю
Istm_model	5000	100	[32, True, 100, 0.2, 0.2, 1, 128, 3]	5ec90a460dd9443be84e9fe5

Моделі Згорткової нейронної мережі

Назва моделі	Кількість найчастіших слів	Довжина тексту	Конфігурація моделі	Назва файлу з моделлю
cnn_model	5000	100	[32, 3, 32, 2, 250, 1, 128, 3]	5ec909970dd9443be84e9fc8

Моделі Комбінованої нейронної мережі

Назва моделі	Кількість найчастіших слів	Довжина тексту	Конфігурація моделі	Назва файлу з моделлю
cnn_lstm_model	5000	100	[32, 3, 32, 2, 100, 32, 5, 0.05, 1, 1]	5ec90976dd9443be84e9fbb

Провести аналіз тексту

Вигляд веб-сторінки авторизації

Я ↻  127.0.0.1:5000

Сторінка авторизації

Введіть, будь ласка, свій логін та пароль:

Логін: Yevhen

Пароль: Ввійти

Новий користувач? Натисніть сюди

Вигляд веб-сторінки реєстрації

Реєстрація нового користувача

Введіть логін:

Введіть пароль:

Повторіть пароль:

Створити користувача

Вигляд веб-сторінки адміністратора системи

Користувачі системи.

Номер користувача,№	Логін користувача	Доступ до моделей	Змінити доступ
5ec842ec8d4e6cd2c158480d	Yevhen	Hi	Змінити доступ
5ec90b0238518da602d7fe0a	Yura	Так	Змінити доступ

					ДП 6303.07.000 КЕ					
Зм.	Арк.	№ документа	Підпис	Дата	Креслення вигляду екранних форм	Літера			Маса	Масштаб
Розробив		Блінков Є.М.								
Перевірів		Жданова О.Г.								
Т. кон.						Аркуш 1			Аркушів 2	
					Інформаційна технологія визначення тональності текстів	КПІ ім. Ігоря Сікорського кафедра АСОІУ гр. ІС-63				
Н. кон.		Проскура С.Л.								
Затвердив		Жданова О.Г.								

Таблиця з результатами роботи алгоритмів

Результати роботи моделей визначення тональності тексту:

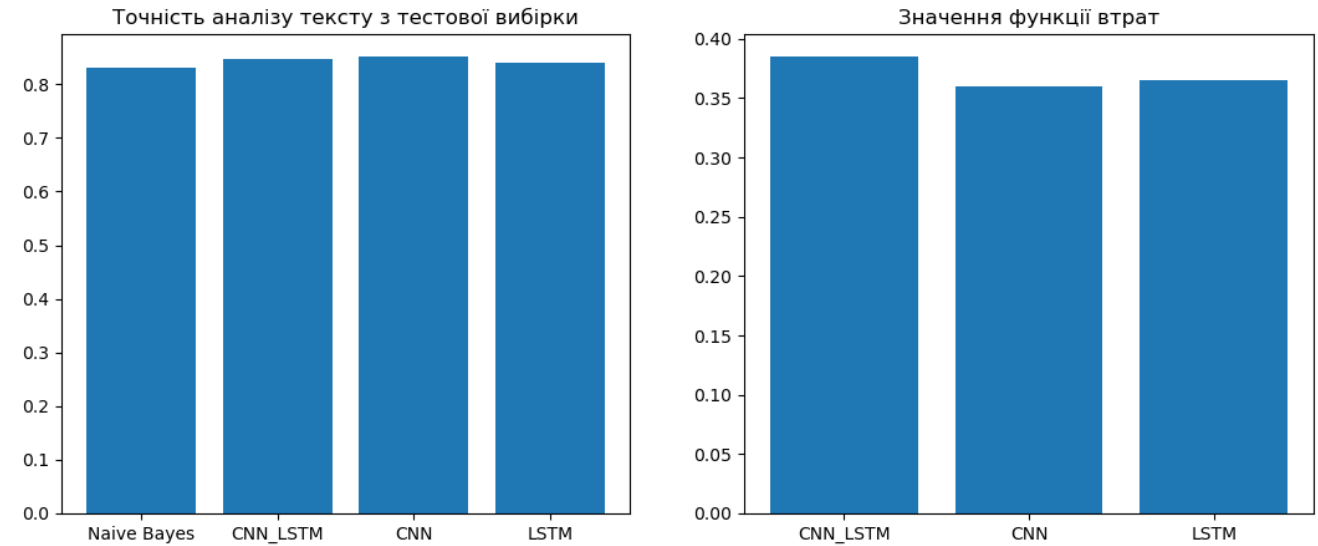
Введений користувачем текст:The movie was awesome. I love it

Назва моделі	Значення тональності	Визначений клас тексту
Рекурентна нейронна мережа з довгою короткочасною пам'яттю (ДКЧП)	0.699	Позитивний
Згорткова нейронна мережа	0.487	Негативний
Комбінована нейронна мережа(ДКЧП+ЗНМ)	0.822	Позитивний

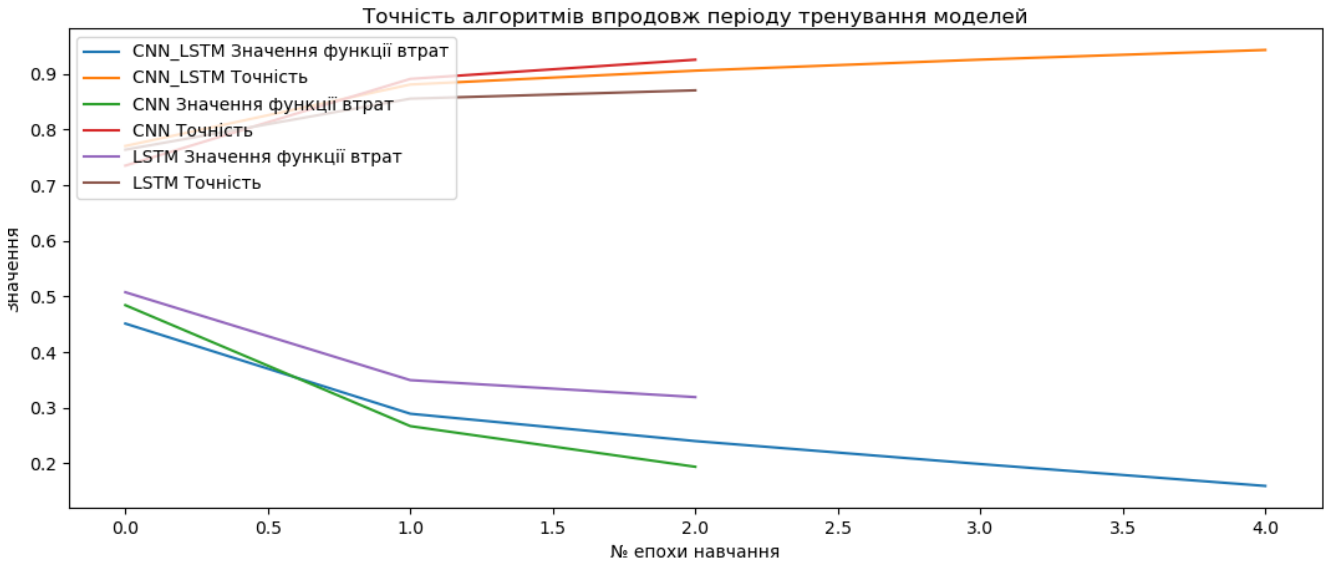
Фрагмент таблиці зі статистикою навчання та тестування моделей

Згорткова нейронна мережа	Назва етапу	Статистика моделі			21.682902812957764
	Навчання	Номер епохи	Точність моделі	Значення функції втрат	
		0	0.73516	0.4842801664113998	
		1	0.89108	0.26686837924003604	
		2	0.92556	0.19378582264900207	
	Тестування	Номер епохи	Точність моделі	Значення функції втрат	
		0	0.85092	0.3598198759508133	
	Навчання	Номер епохи	Точність моделі	Значення функції втрат	138.97327184677124
		0	0.7638	0.5075968775939942	
		1	0.85552	0.34953057083129885	
		2	0.87048	0.3190530140590668	
Рекурентна нейронна мережа з ДКЧП	Тестування	Номер епохи	Точність моделі	Значення функції втрат	
		0	0.84116	0.3648605983018875	

Гістограми точності класифікації текстів з тестової вибірки



Порівняльний графік збільшення точності та зниження значення функції втрат моделей нейронних мереж впродовж тренування



					ДП 6303.07.000 КЕ						
						Креслення вигляду екранних форм	Літера			Маса	Масштаб
Зм.	Арк.	№ документа	Підпис	Дата							
Розробив	Блінков Є.М.										
Перевірив	Жданова О.Г.										
Т. кон.							Аркуш 2			Аркушів 2	
					Інформаційна технологія визначення тональності текстів	КПІ ім. Ігоря Сікорського кафедра АСОІУ гр. ІС-63					
Н. кон.	Проскура С.Л.										
Затвердив	Жданова О.Г.										