

## ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ АНАЛІЗУ СТАТИСТИЧНИХ ПРОФІЛІВ ТЕКСТІВ

Гулик Юрій Володимирович

аспірант кафедри прикладної лінгвістики

Національного університету «Львівська Політехніка»

вул. Степана Бандери, 30, Львів, Україна

*У статті проаналізовано можливості використання штучного інтелекту для формування статистичних профілів текстів на основі аналізу творів видатного українського митця, Андрея Шептицького. Метою дослідження є вивчення особливостей пастирських листів Шептицького, як лінгвістичного явища, в його синтаксичній, морфологічній унікальності й оригінальності в контексті використання штучного інтелекту. Частота вживання різних частин мови, таких як іменники, дієслова, прийменники та ін., співвідношення кількості абзаців, речень та їх синтаксичного характеру у творах митця вимагає нестандартного підходу до аналізу його послань. Результати дослідження підкреслюють важливість іменників та дієслів у формуванні смислово-синтаксичних конструкцій у текстах Шептицького, що є новизною у вивченні його творчості. Стаття пропонує нові можливості використання штучного інтелекту для аналізу лінгвістичних текстів, спрямованих на розуміння мовленнєвої структури та творчої мови письменника. Таким чином, дослідження спрямоване на розширення знань у галузі мовознавства за допомогою інструментів штучного інтелекту та статистичного аналізу текстів. Використання програмного забезпечення з елементами штучного інтелекту виявляється особливо цінним у контексті аналізу текстів великого обсягу, дозволяючи виявляти їхні унікальні риси. Зазначений підхід не лише дозволяє ефективно визначати частотність лексичних одиниць, але і проводити стилістичний та синтаксичний аналіз, а також встановлювати семантичні відтінки використаних слів і виразів. Програмні рішення з елементами штучного інтелекту використовують методи обробки природної мови, що сприяє структуруванню, аналізу та узагальненню великих обсягів текстуальної інформації. Ці інструменти стають не тільки допоміжним елементом для дослідників у лінгвістиці, а й важливим кроком у розвитку комп'ютерних технологій, сприяючи глибокому розумінню текстів та їхніх структурних особливостей. Програмні рішення, які базуються на методах обробки природної мови, не лише допомагають виявляти ключові особливості синтаксичної та морфологічної структури текстів, але й відкривають нові можливості для розуміння творчого внеску митця. Відповідно, взаємодія мовознавства та сучасних технологій надає нові можливості для розвитку інтелектуального аналізу текстів та вивчення мовної естетики в контексті штучного інтелекту.*

**Ключові слова:** штучний інтелект, статистичний профіль, текстовий аналіз, мовленнєва структура.

### Постановка проблеми в загальному вигляді та обґрунтування її актуальності.

Останні роки характеризуються зростаючим інтересом до використання штучного інтелекту (ШІ) у різних сферах, зокрема для аналізу текстових даних. Однією з таких сфер є створення статистичних профілів текстів, що надають цінну інформацію про їхній зміст і характеристики. У цій статті ми дослідимо методи та перспективи використання штучного інтелекту для створення таких профілів (Levchuk, A. 2015).

Одним з ключових способів створення статистичних профілів текстів є обробка природної мови (Natural Language Processing, NLP). NLP використовує алгоритми та обчислювальні моделі для аналізу та розуміння людської мови. Застосовуючи методи NLP до великого корпусу текстів, можна виділити різні статистичні характеристики, такі як частота слів, довжина

речень і синтаксичні патерни. Ці особливості можна використовувати для створення профілів, які фіксують унікальні характеристики різних текстів.

Інший спосіб, який може бути використаний, – це машинне навчання. Алгоритми машинного навчання можна тренувати для набору маркованих текстів, щоб вивчити закономірності та взаємозв'язки між різними ознаками. Після навчання ці алгоритми можна використовувати для класифікації нових текстів і створення статистичних профілів на основі вивчених закономірностей. Такий підхід дозволяє створювати профілі, які є не лише описовими, але й прогностичними, оскільки алгоритми ML можуть робити висновки про зміст і стиль текстів (Levchuk, A. 2015).

Технології використання ШІ для створення статистичних профілів текстів досить різно-

манітні. Такі профілі можуть бути використані в широкому спектрі застосувань, включаючи виявлення плагіату, визначення авторства, аналіз настроїв і рекомендацію контенту. Наприклад, порівнюючи статистичні профілі різних текстів, можна виявити випадки плагіату або визначити авторство анонімних текстів. Аналогічно, аналізуючи настрої, виражені в текстах, можна оцінити громадську думку або спрогнозувати поведінку споживачів.

Однак існують також виклики та обмеження, пов'язані з використанням ШІ для створення статистичних профілів текстів. Однією з проблем є потреба у великих і різноманітних навчальних наборах даних. Алгоритми ML вимагають значної кількості маркованих даних для вивчення значущих закономірностей, а отримання таких наборів даних може бути трудомістким і дорогим. Крім того, точність і надійність згенерованих профілів залежать від якості навчальних даних і продуктивності алгоритмів ШІ.

Дослідження залучення штучного інтелекту (ШІ) для створення статистичних профілів текстів є актуальною проблемою у сучасній науці. Основною складністю полягає у визначенні самого інтелекту, що включає вміння адаптуватися до нових ситуацій, розв'язувати проблеми, відповідати на запитання та формувати плани. Розрізнити інтелект людини від інтелекту дельфінів чи людиноподібних мавп є значною складністю в цьому контексті. Штучний інтелект, у свою чергу, представляє собою науку про системи, які здатні виконувати дії, що сприймаються спостерігачем як розумні. Це означає використання методів, побудованих на основі розумової поведінки людей і тварин, для розв'язання складних проблем. Хоча теорія штучного інтелекту є новою галуззю інтелектуальних досліджень, її коріння сягає глибоко в минуле. Розуміння ШІ вимагає глибоких знань у різних галузях, зокрема лінгвістики (Levchuk, A. 2015). Лінгвістика, як галузь, вносить вагомий внесок у розуміння та розвиток штучного інтелекту, оскільки вона досліджує мовні структури та їхню взаємодію, що є ключовим для розвитку алгоритмів обробки природної мови в рамках штучного інтелекту.

**Аналіз останніх досліджень.** Останні дослідження та публікації у сфері використання штучного інтелекту для створення статистичних профілів текстів підтверджують значний прогрес у цій галузі. Ряд визначних світових та вітчизняних вчених, таких як В. Маккалок, В. Піттс, Ф. Розенблатт, К. Цузе, В. Шикард, М. Бондаренко, Т. Бровченко, Т. Вінцюк, М. Деркач, О. Карпов, Е. Нушякін та інші, від-

дали свою увагу розвитку штучного інтелекту (Sosnina, 2013).

У цій галузі спостерігається широкий спектр досягнень. Нові алгоритми обробки природної мови, базуючись на машинному навчанні, дозволяють відображати складні структури текстів. Технології глибокого навчання спрощують аналіз та відображення смислових взаємозв'язків між словами і фразами.

Дослідницькі зусилля спрямовані на створення більш точних та ефективних алгоритмів для виявлення авторства, визначення стилю та семантичного аналізу текстів. При цьому збільшується увага до питань етики та безпеки, особливо у використанні штучного інтелекту для аналізу та класифікації конфіденційних даних.

Важливим вдосконаленням є розробка технік оцінки точності та достовірності отриманих статистичних профілів. Це сприяє підвищенню якості аналізу та довіри до використання штучного інтелекту в даній області.

Основні напрямки досліджень таких вчених, як В. Шикард, М. Бондаренко, Т. Бровченко, Т. Вінцюк, М. Деркач, О. Карпов вчених стосуються використання машинного навчання, обробки природної мови та глибокого навчання для аналізу текстів, створення статистичних профілів та моделювання мовних паттернів. Роботи цих дослідників відображають значний внесок у розуміння та розвиток технологій, що використовують штучний інтелект для аналізу текстів.

Проте, попри значний прогрес у цій галузі, залишаються невирішеними певні аспекти. Наприклад, покращення точності аналізу текстів у різних мовах залишається актуальною проблемою. Розробка більш складних моделей для виявлення контекстуальних відмінностей у текстах різних стилів та жанрів також є важливим завданням. Крім того, оптимізація алгоритмів для роботи з великими обсягами текстової інформації залишається однією з ключових викликів у цій галузі.

**Мета** цього дослідження полягає у використанні штучного інтелекту для створення статистичних профілів лінгвістичних текстів, з урахуванням важливих аспектів, які розширюють наявні підходи у цій області.

Ця публікація спрямована на введення нових наукових фактів, підкріплених аналізом останніх досліджень та публікацій, що є фундаментом для розвитку та уточнення існуючих концепцій. Вона базується на припущенні, що використання специфічних моделей глибокого навчання та вдосконалення алгоритмів обробки природної мови може ефективно під-

вищити точність та рівень аналізу текстів для створення статистичних профілів.

**Виклад основного матеріалу дослідження.** Дослідження застосування штучного інтелекту (ШІ) для створення статистичних профілів текстів привертає увагу до важливих аспектів взаємодії між ШІ та розумовою діяльністю людини та пропонує спробу перенести особливості головного мозку на штучні інтелектуальні системи, зокрема – паралельну обробку великих обсягів різноманітної інформації (Levchuk, A. 2015).

Інтенсивні дослідження націлені на виявлення ключових різниць у способі передачі та обробки інформації між людським мозком та інтелектуальними машинами. Мозок, здатний до одночасного функціонування мільйонів нейронів, забезпечує паралельну обробку різноманітної інформації. Це спонукало вчених використовувати нанотехнології для створення прототипу ШІ, аналогічного до клітин головного мозку, що здатний одночасно зчитувати та обробляти інформацію в різних напрямках.

Значну увагу також приділяють можливостям, які відкриваються для машин зі штучним інтелектом у сфері отримання та обробки інформації. Як і люди, ці системи можуть використовувати різні сенсори, здійснювати сприйняття через слух та зір, а також користуватися електронними носіями, наприклад, Інтернетом. Ця здатність до обробки та зберігання інформації безпосередньо впливає на їхню ефективність (Anant, 2023).

**Результати дослідження** статистичного профілю Пастирських послань Андрея Шептицького базувалися на використанні методів штучного інтелекту. Важливо розуміти, що штучний інтелект не є повністю ідентичним людському. Мозок людини працює більш ефективно та швидше без складних алгоритмів, які потрібні для інтелектуальних машин. Таким чином, ШІ не створює точний аналог людського інтелекту, але спрямовується на досягнення подібних результатів за допомогою нових підходів (Anant, 2023).

Аналіз пастирських послань Андрея Шептицького було сформовано за допомогою програмної платформи Python. Використання цієї програми дозволило провести глибокий аналіз текстів з використанням широкого спектру інструментів та алгоритмів, що допомогли виявити основні стилістичні та лінгвістичні особливості послань. Python, як потужний інструмент програмування, забезпечив можливість автоматизації процесів обробки великої кількості текстових даних та витягнення

ключової інформації для подальшого аналізу. Такий підхід дозволяє вченим і дослідникам зосередитися на глибокому розумінні текстів і виявленні нюансів, що важливі для розуміння специфіки вираження думок та ідей у відповідних контекстах.

Напрямки подальших досліджень у цій області мають потенціал для розвитку більш ефективних алгоритмів паралельної обробки даних, адаптації до широкого спектру сенсорів для отримання інформації та постійного удосконалення систем навчання машин. Це сприятиме підвищенню їхньої продуктивності та точності у завданнях, що потребують інтелектуальних здібностей.

Штучний інтелект знаходиться на перехресті свого розвитку, де періоди «злету» і «падіння» відображають його молоді природу. На сьогоднішній день галузь штучного інтелекту перебуває у фазі «спаду», використовуючи досягнуті результати в різних сферах, що відображається у високій якості досліджень та аналізу статистичного профілю (Sisyak, 2016).

Результати аналізу текстів виявили значну різноманітність у використанні абзаців та їхній структурі. Програма Rayton допомогла підрахувати загальну кількість абзаців у 52 документах, показавши великі відмінності між максимальними та мінімальними значеннями. Наприклад, середня кількість абзаців у документі складає близько 127, що свідчить про варіативність у використанні мовленнєвих одиниць у текстах.

Детальний аналіз розкрив особливості структури текстів: використання рубрикації, поділ на підрозділи з власними заголовками та параграфами. Така організація структури дозволяє створювати складні синтаксичні цілі, об'єднуючи кілька абзаців у межах одного параграфа. Абзаци у документі виявилися короткими, більшість складалася з 1-3 речень, що полегшує сприйняття інформації та засвоєння причинно-наслідкових зв'язків у тексті. Це важливо для зручності сприйняття та розуміння інформації, сприяючи кращому засвоєнню контекстуальних зв'язків у тексті.

Документ з найменшою кількістю абзаців виділяється своєрідним підходом до структурування. Тут автор не застосовує рубрикацію та поділ на підрозділи, а замість цього поділяється своєю думкою та побажаннями. Такий підхід є прикладом авторського стилю та відображає його відношення до тексту.

Загальний аналіз показав, що немає прямої кореляції між обсягом документу та кількістю абзаців, оскільки довжина абзаців та їхня структура різняться в кожному документі.

**Результати дослідження** методів та перспектив застосування штучного інтелекту для створення статистичних профілів текстів на основі Статистичного профілю Пастирських послань Андрея Шептицького показали варіативність у структурі абзаців між різними документами.

Детальний аналіз виявив, що навіть при схожих кількісних показниках абзаців у документах, кожен абзац має власну структуру та форму. Найкоротший абзац містив лише 22 слова, що різнилося від найдовшого, що включав у себе 96 слів. Це свідчить про велику різноманітність у використанні мовленнєвих одиниць та підходів до композиції текстів.

Аналіз статистичних даних з текстів Андрея Шептицького вказує на те, що більшість абзаців (97.77%) складається з невеликої кількості речень: одного (43.58%), двох (23.01%) та трьох (13.27%). Значні відмінності відзначаються в тому, що у деяких випадках Андрей Шептицький використовував абзаци із великою кількістю речень, навіть до 50 у одному абзаці.

Цікаво врахувати, що складність тексту може залежати від ступеня агрегації мовних одиниць, наприклад, літер у словах чи слів у реченнях. Велика кількість елементів утворюється до певного об'єкта, ускладнюючи сприйняття читачем взаємозв'язків між ними. Слова коротшої довжини, як правило, використовуються частіше, тоді як довші слова можуть бути менш зрозумілими.

Аналогічно, довжина речень має вплив на їх зрозумілість. Довші речення з більшою кількістю довгих слів можуть бути складнішими для сприйняття. Це підтверджується аналізом речень у тексті Пастирських послань Андрея Шептицького. Чим коротше речення та менше в них довгих слів, тим зрозуміліше текст.

Виявлення довжини речень стало інструментом для розкриття індивідуального стилю. Багато дослідників погоджуються, що довжина речення не може бути необмеженою, хоча існують винятки. Розподіл вживання слів у реченнях показує, що більше за всі інші вживаються речення довжиною від 10 до 15 символів, що становить майже чверть усіх речень.

Деякі речення у тексті містять більше 60 слів, але це становить невелику частину. Використання занадто довгих речень може викликати «розрідженість» інформації та «водянистість» викладу, що порушує ритм сприйняття і утруднює засвоєння думки автора.

У текстах можуть виявлятися певні стилістичні особливості, такі як лексичне і граматичне повторення, тавтологія та надмірні

пояснення, що утворюють інформаційний надлишок. Ці особливості можуть бути використані для аналізу текстів.

Статистичний підхід у сполученні зі стилістикою може відображати тенденцію визначення довжини речення як частини загального прагнення до статистичного вивчення всіх аспектів мови. Такі аналізи можуть розкривати різницю у довжині речень залежно від жанрових характеристик тексту, а також у розмірах складнопідрядних і простих речень (Luo, 2023).

Для порівняння використано середні величини довжини речень у творах різних жанрів. Наприклад, драматургія відрізняється короткими реченнями — майже половина (49.73%) містить не більше 3 слів. Подальший аналіз показує, що чим більше слів у реченні, тим менш ймовірно зустріти його у драматичних творах.

Отримані результати аналізу вказують на суттєві схожості між дослідженими текстами Пастирських послань Андрея Шептицького та прозою української мови. Особливу увагу привертає достатньо велика кількість довгих речень у текстах Шептицького, що створює специфічну особливість його творчого стилю. Це може бути важливим елементом в розумінні його манери висловлення та впливу на сприйняття тексту читачами.

Для аналізу частотних параметрів частин мови, зокрема слів, використовувалася лематизація. Цей процес полягає у виведенні лексикографічних форм слів на основі їх словоформ. Лема відображає вихідну форму слова, тоді як токен – кожна конкретна словесна одиниця у тексті. Наприклад, для слова «я» у реченні «*І тоді я оповім їй про те, що зі мною вчинив цей чаклунський вечір*» лемою є «я», а слововживанням – «зі мною» (Sheptytsky, 2007, с. 456).

Під час лематизації текстів визначали леми словоформ, їхню частоту вживання і розподіл за частинами мови для подальшого детального аналізу.

Отримані результати аналізу частотних параметрів текстів Пастирських послань Андрея Шептицького розкривають важливі аспекти структури мови та слововживання в цих творах.

Співвідношення частин мови в тексті вказує на присутність особливого стилістичного підходу автора. Наприклад, у цьому випадку іменники використовуються частіше за інші частини мови, відіграючи значну роль у структурі та смисловій спрямованості тексту. Їхня вагова частка (більше 37%) у вживанні частин мови відображає нахил Шептицького до деталізації і побудови образів, розширення змісту текстів.

Дієслова, що становлять майже 29% від усіх частин мови, є ще однією ключовою складовою структури. Це вказує на активність дії та емоційну насиченість тексту, а також на те, що автор акцентує увагу на використанні різних дієслівних форм для розкриття своїх думок та ідей.

Варіативність вживання частин мови у тексті є складним явищем, яке залежить від багатьох чинників, включаючи стиль, жанр та індивідуальні особливості автора. Цей аналіз сприяє кращому розумінню структури та особливостей текстів, допомагаючи виявити і висвітлити унікальні аспекти мовної естетики в творчості Шептицького.

Лематизація, як процес, зберігає основний контекст слів, розкриваючи відтінки вживання та їхніх лексичних форм у тексті. Наприклад, для аналізу текстів Пастирських послань Андрея Шептицького застосовувалася лематизація для виведення лексикографічних форм слів на основі їх словоформ. Цей підхід дозволив визначити леми словоформ, їхню частоту вживання та розподіл за частинами мови для подальшого детального аналізу.

Отримані результати розкривають важливі аспекти структури мови та слововживання в творах Шептицького. Наприклад, співвідношення частин мови в тексті вказує на присутність особливого стилістичного підходу автора. Іменники використовувалися частіше за інші частини мови, що відображало нахил Шептицького до деталізації, розширення змісту текстів.

Дієслова, що становлять значну частку, вказують на активність дії та емоційну насиченість тексту. Вони допомагають автору акцентувати увагу на різноманітних дієслівних формах для розкриття своїх думок та ідей.

Варіативність вживання частин мови у тексті є складним явищем, що відбиває стиль, жанр та індивідуальні особливості автора. Цей аналіз допомагає краще розуміти структуру та особливості текстів, виявляючи унікальні аспекти мовної естетики в творчості Шептицького.

На основі нашого аналізу виявлено, що іменники, як ключовий елемент мовленнєвої структури, не лише виступають у ролі підмета у реченнях, але й часто приймають функції додатка. Це підтверджує повноту синтаксичної природи речень, які створив митрополит Шептицький у своїх посланнях. Дослідження синтаксичної структури текстів підкреслило важливість ретельного вивчення семантики іменникових форм у побудові змісту і тонкощів передачі ідей.

Під час аналізу структури текстів також виявлено, що довжина документу не завжди

має прямий зв'язок з кількістю абзаців чи речень. Проте помітна тенденція до більш дрібного членування абзаців, складення їх з меншою кількістю речень, що є характерною особливістю будови текстів.

Загальною висновком нашого дослідження є підкреслення важливості статистичного аналізу для розуміння структури та мовленнєвого виразу в творчості Андрея Шептицького. Це також підкреслює значущість ролі іменників та дієслів у формуванні змісту та виразності його текстів. Це дозволяє краще розкрити та зрозуміти глибину і майстерність виразного мовлення, яке використовував митрополит Шептицький у своїй творчості.

Результати нашого дослідження свідчать про великий потенціал, який забезпечує застосування штучного інтелекту для аналізу мовленнєвого виразу. Використання таких методів дозволяє виявити ключові особливості мови та формувати статистичні дані, які розкривають різноманітні стилі, авторські риси та синтаксичні особливості текстів.

Аналіз на значенні іменників та дієслів у вираженні змісту та виразності текстів. Іменники є основною частиною мови, що називає предмети зображення та слугує основою для побудови тексту. Дієслова відображають активність дії та емоційну насиченість тексту, сприяючи розкриттю авторських думок та ідей.

Дослідження синтаксичної структури текстів підтверджує широкий спектр використання іменників, які часто не лише виступають у ролі підмета, а й у функції додатка. Це свідчить про їхню важливу синтаксичну роль у формуванні повних синтаксичних структур речень.

Крім того, аналіз показав, що довжина тексту не завжди прямо пов'язана з кількістю абзаців чи речень. Тим не менше, спостерігається тенденція до більш дрібного членування абзаців, що є важливою рисою структури текстів.

**Висновки.** В цілому, наше дослідження демонструє важливість та значущість статистичного аналізу для розуміння структури та особливостей мовленнєвого виразу у вивченні текстів. Використання штучного інтелекту для формування статистичних профілів текстів відкриває нові можливості для більш глибокого та об'єктивного аналізу мовленнєвих особливостей у літературі та письмових текстах.

Отже, методології та перспективи використання методів штучного інтелекту для створення статистичних профілів текстів є дієвими. Алгоритми NLP та ML можуть бути застосовані для вилучення та аналізу різноманітних ознак текстів, що призводить до генерації описових

та предикативних профілів. Однак для забезпечення надійності та корисності згенерованих профілів необхідно вирішити проблеми, такі як доступність навчальних даних і точність алгоритмів. З подальшим розвитком технологій штучного інтелекту та наявністю великих масивів даних створення статистичних профілів текстів має потенціал для революційних змін у різних сферах – від академічних до промислових.

Аналіз Пастирських послань Андрея Шептицького був проведений за допомогою програми Python та інших електронних інструментів, які використовують штучний інтелект для систематизації лексики та аналізу мовленнєвих особливостей. Цей підхід дозволяє створювати статистичні профілі текстів, які відображають ключові моменти та структурні особливості творчого доробку Шептицького.

Застосування програмного забезпечення з елементами штучного інтелекту украй цінне для аналізу текстів великого обсягу та виявлення їхніх унікальних рис. Використання цих інструментів дозволяє ефективно виявляти частотність лексичних одиниць, проводити стилістичний та синтаксичний аналіз, а також встановлювати семантичні відтінки використаних слів і виразів.

Програмні рішення з елементами штучного інтелекту у своїй роботі використовують методи обробки природної мови, що дозволяє структурувати, аналізувати та узагальнювати великі обсяги текстуальної інформації. Використання цих інструментів стає не тільки допоміжним елементом для дослідників у лінгвістиці, а й важливим кроком у розвитку комп'ютерних технологій, які сприяють глибокому розумінню текстів та їхніх структурних особливостей.

#### ЛІТЕРАТУРА

1. Левчук А. Штучний інтелект: лінгвістичні аспекти. URL: [http://www.irbis-nbuv.gov.ua/cgi-bin/irbis\\_nbuv/cgiirbis\\_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP\\_meta&C21COM=S&2\\_S21P03=FILEA=&2\\_S21STR=Nvvnuffm\\_2015\\_4\\_41](http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&C21COM=S&2_S21P03=FILEA=&2_S21STR=Nvvnuffm_2015_4_41).
2. Перебийніс В. С. Статистичні параметри стилів. Київ.: Наукова думка, 1967. 260 с.
3. Сісяк П. Штучний інтелект – революція, надія чи утопія? URL: <https://www.imena.ua/blog/ai-revolution/>.
4. Сосніна А. Штучний інтелект як наука та технологія створення інтелектуальних роботів. URL: <https://naub.oa.edu.ua/shtuchnyj-intelekt-yak-nauka-ta-tehnolohiya-stvorennya-intelektualnyh-robotiv/>.
5. Шептицький Андрей. Пастирські послання 1899-1914 рр., т. 1. Львів: Видавництво «АРТОС» 2007. 1014 с.
6. Anant. What is Linguistics? The future of AI and Linguistics. *Medium*. URL: <https://medium.com/@anant3104/what-is-linguistics-the-future-of-ai-and-linguistics-6f0f9117a7bd>.
7. Luo, J. Three Innovative uses of Artificial Intelligence for Languages: Understanding, Preserving and Learning | Text Inspector. URL: <https://textinspector.com/artificial-intelligence-in-languages/>.

#### REFERENCES

1. Levchuk, A. (2015). Shtuchnyi intelekt: lnhvistychni aspekty [Artificial intelligence: Linguistic aspects]. *Scientific Bulletin of Lesya Ukrainka Eastern European National University*. <https://rgf.vnu.edu.ua/index.php/rgf/article/view/251/250> [in Ukrainian]
2. Perebyinis, V. S. (1967). Statystychni parametry styliv [Statistical parameters of styles]. *Scientific thought*. [in Ukrainian]
3. Sisyak, P. (2016). Shtuchnyi intelekt – revoliutsiia, nadiia chy utopiia?[Artificial intelligence – revolution, hope or utopia?] <https://www.imena.ua/blog/ai-revolution/>. [in Ukrainian]
4. Sosnina, A. (2013). Shtuchnyi intelekt yak nauka ta tekhnolohiia stvorennia intelektualnykh robotiv [Artificial intelligence as a science and technology of creating intelligent robots]. <https://naub.oa.edu.ua/shtuchnyj-intelekt-yak-nauka-ta-tehnolohiya-stvorennya-intelektualnyh-robotiv/> [in Ukrainian]
5. Sheptytsky, A. (2007). Pastyrski poslannia 1899-1914 [Pastoral epistles of 1899-1914]. *ARTOS Publishing House*. [in Ukrainian]
6. Anant. (2023). *What is Linguistics? The future of AI and Linguistics*. Medium. <https://medium.com/@anant3104/what-is-linguistics-the-future-of-ai-and-linguistics-6f0f9117a7bd>
7. Luo, J. (2023). *Three Innovative uses of Artificial Intelligence for Languages: Understanding, Preserving and Learning | Text Inspector*. Text Inspector. <https://textinspector.com/artificial-intelligence-in-languages/>

## USING ARTIFICIAL INTELLIGENCE TO ANALYZE STATISTICAL PROFILES OF TEXTS

**Hulyk Yuriy Volodymyrovych**

*Postgraduate Student at the Department of Applied Linguistics*

*Lviv Polytechnic National University*

*30, Stepan Bandera str., Lviv, Ukraine*

*The article analyses the possibilities of using artificial intelligence to generate statistical profiles of texts based on the analysis of the works of the prominent Ukrainian artist Andrey Sheptytsky. The aim of the study is to examine the peculiarities of Sheptytsky's pastoral letters as a linguistic phenomenon, in its syntactic, morphological uniqueness and originality in the context of the use of artificial intelligence. The frequency of use of different parts of speech, such as nouns, verbs, prepositions, etc., the ratio of the number of paragraphs, sentences and their syntactic character in the works of the artist requires a non-standard approach to the analysis of his messages. The results of the study emphasise the importance of nouns and verbs in the formation of semantic and syntactic constructions in Sheptytsky's texts, which is a novelty in the study of his work. The article offers new possibilities of using artificial intelligence to analyse linguistic texts aimed at understanding the linguistic structure and creative language of the writer. Thus, the study is aimed at expanding knowledge in the field of linguistics with the help of artificial intelligence tools and statistical text analysis. The use of software with elements of artificial intelligence is particularly valuable in the context of analysing large texts, allowing to identify their unique features. This approach not only makes it possible to effectively determine the frequency of lexical items, but also to conduct stylistic and syntactic analysis, as well as to establish the semantic connotations of the words and expressions used. Software solutions with artificial intelligence elements use natural language processing methods to help structure, analyse and summarize large amounts of textual information. These tools are becoming not only an auxiliary element for linguistic researchers, but also an important step in the development of computer technology, contributing to a deeper understanding of texts and their structural features. Software solutions based on natural language processing methods not only help to identify key features of the syntactic and morphological structure of texts, but also open up new opportunities for understanding the creative contribution of an artist. Accordingly, the interaction of linguistics and modern technologies provides new opportunities for the development of intellectual text analysis and the study of linguistic aesthetics in the context of artificial intelligence.*

**Key words:** *artificial intelligence, statistical profile, text analysis, speech structure.*