

DOI: <https://doi.org/10.36910/6775-2524-0560-2021-45-09>

УДК 004.912

Гарбузенко Олександр Віталійович, магістр

П'ятикоп Олена Євгенівна, к.т.н., доцент

<https://orcid.org/0000-0002-7731-3051>

ДВНЗ «Приазовський державний технічний університет», м. Маріуполь, Україна

## ПОРІВНЯННЯ ВЛАСТИВОСТІ ЗАСВОЄННЯ СЕМАНТИЧНИХ ЗВ'ЯЗКІВ МІЖ СЛОВАМИ ПРИРОДНОЇ МОВИ МОДЕЛЯМИ МЕТОДУ WORD2VEC У ЗАДАЧІ АНАЛІЗУ НАСТРОЇВ

Гарбузенко О. В. П'ятикоп О. Є. Порівняння властивості засвоєння семантичних зв'язків між словами природної мови моделями методу Word2Vec у задачі аналізу настроїв. Дана робота присвячена дослідженню ефективного визначення настрою англійськомовних постів з соціальних мереж, що базується на перетворенні слів у векторні представлення за допомогою методу Word2Vec. У роботі описані та проаналізовані існуючі методи сентимент аналізу, проаналізовано моделі Continuous Bag of Words (CBOW) та Skip-gram у складі методу Word2Vec, проведено порівняння їх властивостей при засвоєнні семантичних зв'язків між словами природної мови. Описано експериментальне дослідження щодо використання зазначених моделей при різних функціях тренування.

**Ключові слова:** сентимент аналіз, визначення настроїв, англійськомовний пост, векторне представлення, машинне навчання, skip-gram, CBOW.

Гарбузенко А. В. Пятикоп Е. Е. Сравнение свойства усвоения семантических связей между словами естественного языка моделями в составе метода Word2Vec в задаче анализа настроений. Данная работа посвящена исследованию эффективного определения настроения англоязычных постов из социальных сетей, базирующегося на преобразовании слов в векторные представления с помощью метода Word2Vec. В работе описаны и проанализированы существующие методы сентимента анализа, проанализированы модели Continuous Bag of Words (CBOW) и Skip-gram в составе метода Word2Vec, приведено сравнение их свойств при установлении семантических связей между словами естественного языка. Описано экспериментальное исследование использования указанных моделей при различных функциях тренировки.

**Ключевые слова:** сентимент анализ, определение настроений, англоязычный пост, векторное представление, машинное обучение, skip-gram, CBOW.

Harbuzenko O. V. Piatyko E. E. Comparison of the properties of learning semantic relationships between words of a natural language by models of the Word2Vec method in the problem of sentiment analysis. This article is devoted to the research of the effective sentiment analysis method of English-language posts from social networks, based on the conversion of words into vector representations using the Word2Vec method. The paper describes and analyzes the existing methods of sentiment analysis, analyzes the Continuous Bag of Words (CBOW) and Skip-gram models as part of the Word2Vec method, provides a comparison of their properties when establishing semantic relationships between words in a natural language. An experimental study of the use of these models for various training functions is described.

**Keywords:** sentiment analysis, mood determination, social media post, vector representation, machine learning, skip-gram, CBOW.

### Постановка наукової проблеми.

Сьогодні вже важко уявити без можливостей Інтернету та безлічі соціальних мереж, які надають людині можливість обмінюватися думками та емоціями у електронному вигляді. Щодня у соціальних мережах користувачі висловлюють свою точку зору або передбачення про той чи інший об'єкт або подію. Такі записи, частіш за все, мають на меті передавати інформацію та, зокрема, настроїв. Ці відомості представляють собою емоційно забарвлений текст, а множина таких текстів – це актуальне і цінне джерело для соціально-економічних досліджень: від глобальної статистики до думок про певний товар чи послугу. Тому останнім часом швидкий розвиток отримала область аналізу емоційності (sentiment analysis) - сімейство методів обробки природної мови, присвячене ідентифікованню та визначенню емоційного забарвлення тексту.

При цьому, пости з соціальних мереж мають свої особливості – невелике повідомлення може у собі граматичні помилки, неологізми, сарказм та ін. Отже так важливо винаходити та покращувати методи аналізу настроїв для сучасного інформаційного середовища. Актуальним підходом є використання методів машинного навчання, що можуть переймати контекст та зміст сучасної мови у соціальних мережах. При цьому, склад попередньої обробки даних залежить від конкретних обраних методів для визначення настроїв. Як представник методів машинного навчання, у дослідженнях використовується Word2Vec. Цей інструмент, розроблений компанією Google, може навчатися на текстових даних та перетворювати слова у їх векторне представлення, при цьому зберігає семантичні зв'язки всіх слів у певній мірі.

Word2Vec — це назва класу моделей нейронних мереж, які, маючи немаркований навчальний корпус, створюють вектор для кожного слова в корпусі, що кодує його семантичну інформацію. Ці вектори корисні з двох основних причин [6, 10, 11]:

– можна виміряти семантичну подібність між двома словами, обчислюючи косинусну схожість між відповідними векторами слів.

– можна використовувати ці вектори слів як функції для різних контрольованих завдань NLP, таких як класифікація документів, розпізнавання іменованих об'єктів та аналіз настроїв.

Семантична інформація, що міститься в цих векторах, робить їх потужними функціями для виконання цих завдань.[7]

Створені векторні представлення та експертні оцінки настроїв текстів вже використовуються для навчання моделі-класифікатора, що добре працює з векторами характеристик (ознак), за якими слова відносяться до певного класу (у даному випадку, настрою). Зазвичай, це лінійні моделі з різними функціями втрат, наприклад, метод опорних векторів. З точки зору завдання класифікації за настроями, перетворення слів на вектори є способом зниження розмірності та/або стандартизації вихідних даних.

Щоб визначення настроїв англійських постів було ефективним з точки зору правильного визначення класу настрою, необхідно провести експериментальне дослідження, що дозволить виявити, яку з моделей, що представлені методом Word2Vec, доцільніше використовувати для перетворення вихідних текстів у вектори для моделі-класифікатора.

Щоб обрати модель з кращою ефективністю з точки зору кінцевого визначення настроїв тексту, слід вирішити наступні завдання:

– забезпечити моделі навчальною та тестовою вибірками англійських текстів (постів) з соціальних мереж, що мають експертні оцінки (щонайменше – позитивні, негативні чи нейтральний зміст);

– організувати процес попередньої обробки текстів: визначити надлишкові елементи текстів (імена користувачів, посилання, теги, знаки пунктуації, тощо); визначити спосіб попередньої обробки текстів, згідно з NLP підходами, що може покращити навчання семантичним зв'язкам;

– організувати процес навчання моделі перетворення тексту на вектори та визначення якості відображення семантичних зв'язків у векторних представленнях;

– організувати процес навчання класифікатора з оцінкою ефективності класифікації за декількома спеціальними метриками (precision, recall, f1-score та accuracy);

– організувати процес класифікації цільових даних та застосувати методи візуального представлення інформації за результатами класифікації (розподіл екземплярів у класах) для остаточної оцінки ефективності визначення настроїв.

**Метою роботи** є аналіз та порівняння властивостей моделей Continuous Bag of Words (CBOW) та Skip-gram у складі методу Word2Vec при засвоєнні семантичних зв'язків між словами природної мови; вибір моделі з кращим показником для застосування при визначенні (класифікації) настрою англійських постів із соціальних мереж.

#### **Аналіз останніх досліджень і публікацій.**

Як зазначається у багатьох роботах – для вирішення задач сентимент аналізу можна використовувати лексемний метод, метод опорних векторів, наївний класифікатор Баєса та метод максимальної ентропії та ін. Недоліком методу опорних векторів є те, що він здійснює бінарну класифікацію, яка дозволить розділити дані тільки на дві категорії: дані без інформаційно-психологічних впливів та дані з інформаційно-психологічними впливами. Основний недолік наївного класифікатора Баєса - неможливість врахування залежності результату від комбінації ознак (слів). Спільним недоліком лексемного методу, наївного класифікатора Баєса та методу максимальної ентропії є необхідність складання словників, що вимагає тісної співпраці з лінгвістами [1-3].

Для лексемного та деяких інших методів мається складність попереднього етапу сентимент-аналізу: побудова словника та лінгвістична обробка. Але, незважаючи на те, що деякі методи обходяться без попередньої лінгвістичної обробки, точність класифікації може бути нижчою, ніж при зниженні розмірності та приведення вихідних даних до стандартного вигляду. Прикладом стандартизації вихідних даних може бути перетворення слів, речень, або цілих документів у вектори ознак для подальшої класифікації доступними моделями. Це значить, що якість класифікації залежить від вигляду документу, у якому він поданий для обробки та з яких компонентів складається вектор ознак [4, 5].

Було виявлено, що зважування векторів більш ефективне. Це означає, що наявність терміну в тексті важливіше, ніж його частота. Такі вектори представляють собою набір одиниць, на місці, де термін зустрічається у тексті; та нулів, якщо навпаки. Вектори на основі частоти створюються від кількості надходжень певного об'єкту в класі текстів. [7].

Також, є інші підходи, що полягають у створенні, так званих, вкладень слів (англ. word embedding). У цих підходах моделі навчаються на великій кількості вихідних текстів природною мовою, та можуть повертати векторні представлення переданих слів, знаходити схожі слова за допомогою складення та віднімання векторів. При цьому, модель намагається зберегти семантичні зв'язки слів з вихідних даних. До таких методів відносяться набір моделей Word2Vec або GLOVE. Наприклад, Word2Vec складається з двох пов'язаних моделей: Skip-gram та Continuous Bag of Words. Ці моделі використовуються для створення так званих вкладень слів. Skip-gram та Continuous Bag of Words представляють собою плоскі двошарові нейронні мережі. Після навчання вони можуть відтворювати певний лінгвістичний контекст слів, зокрема, семантичні зв'язки різних типів. Для навчання Word2Vec використовує великий набір текстів, наприклад, складом у декілька мільйонів слів. На основі проаналізованих текстів виробляється векторний простір, зазвичай з декількома сотнями вимірів, де для кожного унікального слова з цього навчального набору текстів призначається спеціальний вектор. Дані вектори слів розташовуються в побудованому векторному просторі так, що слова, які мають певні семантичні зв'язки у текстах, представляють схожі вектори, тобто розташовані близько один до одного в побудованому моделями векторному просторі [6, 8]. Використання таких методів у задачі сентимент аналізу є досить новим підходом і має багато аспектів для вивчення. Наприклад, як визначається якість навчання моделі методу Word2Vec і як це впливає на подальшу класифікацію текстів за настроями.

Дослідження методів автоматичного аналізу настроїв в соціальних мережах показало, що найбільш придатними для виявлення у текстах та мультимедійних даних інформаційно-психологічних впливів та емоційної забарвленості є нейронні мережі, оскільки вони не потребують складання словників, обов'язкової попередньої лінгвістичної обробки текстів, можуть застосовуватися до різних типів даних та здатні здійснювати класифікацію за декількома категоріями, що дозволить виявляти різні типи інформаційно-психологічного впливу [2, 3].

Навчання моделей-класифікаторів, заснованих на нейронних мережах, проводиться з вчителем. Рекомендується проводити навчання як експеримент із зворотнім зв'язком та аналізом результатів. Для такого експерименту потрібно визначити систему та її компоненти, у якій проводиться навчання та методи цього навчання. Невід'ємним є визначення підходів до тестування системи, вимірювання характеристик та їх аналіз. Важливим є визначення відношення навчальних та тестових даних. Вдалим рішенням є вибирати для навчання більшу кількість екземплярів (80-90% від загальної кількості), при цьому, випадковим чином вибираючи дані з загального набору на кожній епісі навчання [3, 5].

#### **Виклад основного матеріалу й обґрунтування отриманих результатів дослідження.**

При вивченні підходів до аналізу настроїв було з'ясовано, що методи засновані на лексичному підході (зі словниками), мають слабкості:

- велика залежність від предметної області;
- помилкове визнання нейтрального контексту полярним;
- нездатність обробки сарказму, помилкових слів та неологізмів (до реорганізації).

Отже, доцільним буде використання моделей, що перетворюють слова на вектори, бо вони можуть бути менш залежними від предметної області, а також визначають значущість слів у певних випадках.

З описаних підходів, найновішим є проєкт Word2Vec. При навчанні, може використовуватись одна з двох головних моделей: або skip-gram та CBOW (Continuous Bag of Words), архітектура яких наведена на рисунку 1.

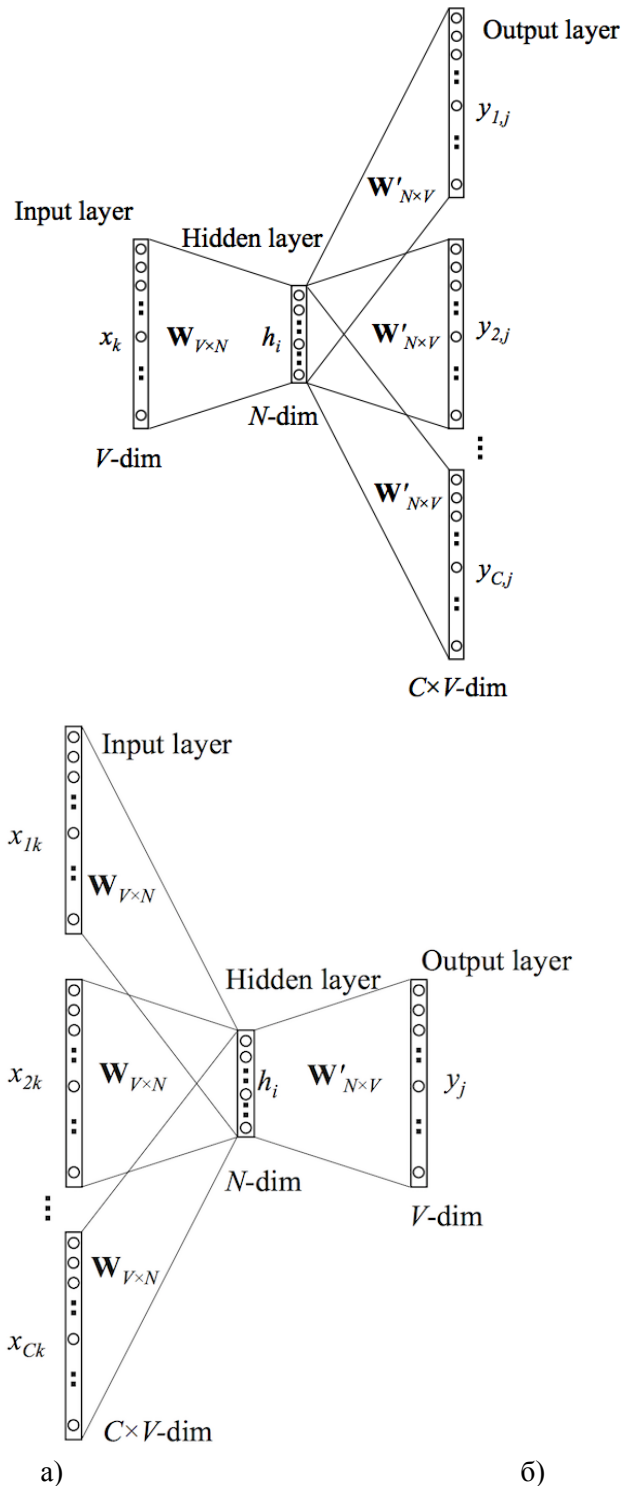


Рисунок 1 – Архітектура нейронних мереж skip-gram [9] та CBOW [10]

Skip-gram – це підклас імовірнісної мовної моделі для передбачення наступного елемента в такій послідовності у вигляді моделі Маркова  $(n - 1)$  порядку. Моделі  $n$ -gram зараз широко використовуються в теорії ймовірності, комунікації, комп'ютерній лінгвістиці (наприклад, статистичній обробці природної мови), обчислювальній біології (наприклад, аналізі біологічної послідовності) та стиснення даних. [13] На рисунку 1 а) представлено схематичне зображення нейронної мережі skip-gram. На представленій схемі  $x$  представляє вектор із одноразовим кодуванням, що відповідає вхідному слову в навчальному екземплярі, а  $\{y_1, \dots, y_C\}$  — вектори, закодовані одноразовим способом, що відповідають вихідним словам у навчальному екземплярі. Матриця  $V \times N$   $W$  є ваговою матрицею між вхідним шаром і прихованим шаром,  $i$ -й рядок якого представляє ваги, що відповідають  $i$ -му слову у словнику. Інтерес представляє навчання вагової матриці  $W$ , оскільки вона містить векторні кодування всіх слів нашого словника (як його рядки).

Кожен вектор вихідного слова також має асоційовану вихідну матрицю  $W' N \times V$ . Також є прихований шар, що складається з  $N$  вузлів (точний розмір  $N$  є навчальним параметром).

На рисунку 1 б) представлено схематичне зображення нейронної мережі Continuous Bag of Words. Зі схеми можна побачити, що модель CBOW є дзеркальним відображенням skip-gram. Вхідний шар складається з одноразово закодованих вхідних контекстних слів  $\{x_1, \dots, x_C\}$  для вікна слів розміру  $C$  і словника розміру  $V$ . Прихований шар є  $N$ -вимірним вектором  $h$ . Нарешті, вихідний рівень є вихідним словом  $u$  у навчальному прикладі, яке також кодується одноразово. Вхідні вектори з'єднані з прихованим шаром за допомогою матриці  $W$  ваги  $V \times N$ , а прихований шар підключений до вихідного шару через матрицю  $W' N \times V$ .

Зазначені моделі, CBOW та skip-gram, видають подібні результати при різних обчислювальних витратах. Але головна особливість цих моделей – здатність уловлювати певні ступені подібності слів: семантичні та синтаксичні закономірності, що однак може бути корисним при аналізі тональності. Деякими дослідженнями було доведено, що модель skip-gram краще навчається семантичним зв'язкам слів, отже, може бути більш доцільною для створення усереднених векторних представлень текстів для подальшої класифікації [4].

Метою навчання моделі skip-gram є пошук уявлень слів, які є корисними для передбачення навколишніх слів у реченні чи документі. Ідея моделі Continuous Bag of Words полягає ж у тому, щоб отримувати визначене за контекстом слово, передаючи набір слів у якості вхідних параметрів.

Модель Word2Vec може бути треновано за допомогою функцій: hierarchical softmax (ієрархічна нормалізована експоненціальна функція) та/або негативним вибиранням (negative sampling) [14-15].

Softmax (нормалізована експоненціальна функція) - це функція вихідного рівня, яка активує кожен із вузлів на останньому етапі обчислення нейронної мережі. Має вигляд, як вираз (1):

$$p(w_j | w_l) = \frac{\exp(v'_{w_j} \cdot v_{w_l})}{\sum_{w=1}^W \exp(v'_w \cdot v_{w_l})} \quad (1)$$

де:  $v_w$  – вхідне векторне представлення слова  $w$ ;  $v'_w$  – вихідне векторне представлення слів;  $W$  – кількість слів у словнику (унікальні слова з навчального набору).

Ієрархічний softmax використовує двійкові дерева, де листя є ймовірністю слів – лист з індексом  $j$  є ймовірністю  $j$ -го слова та має позицію  $j$  у вихідному векторі softmax.

Кожне зі слів може бути досягнуто від кореня до внутрішніх вузлів, які являють собою ймовірнісну масу на цьому шляху.

Ідея використання negative sampling (негативної вибірки) полягає в концепції контрастної оцінки шуму (аналогічно генеративним змагальним мережам), за якою хороша модель має відрізнити підроблений сигнал від реального з допомогою логістичної регресії. Крім того, мотивація негативної мети вибірки аналогічна стохастичному градієнтному спуску: замість того, щоб щоразу змінювати всі ваги з урахуванням усіх тисяч спостережень, які є, використовується тільки  $K$  з них і різко збільшуючи обчислювальну ефективність (залежить від кількості негативних зразків).

Авторами досліджено, що hierarchical softmax дає кращі показники для рідкісних слів, тоді як negative sampling показує кращі показники для слів, що зустрічаються частіше. При цьому, остання функція краще працює з векторами з меншим числом вимірів. Зі збільшенням числа тренувальних епох ієрархічна softmax бути корисною перестає [6, 8, 4]. Проведено дослідження щодо цих тверджень.

Розмір навчального набору текстів та налаштування параметрів моделей може сильно вплинути на якість їх якості. Точність моделей можна покращити наступними способами:

- вибрати більш релевантну, з огляду на навчальні дані та параметри, архітектуру;
- збільшити кількість навчальних текстів;
- збільшенням кількості вимірів для побудови векторних представлень;
- змінити розмір вікна, за яким розглядається контекст слів.
- налаштувати параметри навчання, наприклад, випадкове ігнорування слів з пороговою частотою.

Кожен зі способів може подовжувати час, необхідний для навчання, через збільшення обчислювальної складності задачі [8].

Точність моделей, зазвичай, зростає зі збільшенням кількості навчальних слів, та зі зростанням кількості вимірів векторних представлень. Було досліджено, що збільшення обсягу тренувальних

даних у два рази призводить до зростання обчислювальної складності, що еквівалентно до подвоєння кількості вимірів векторних представлень [6].

При визначенні настроїв для моделей Word2Vec важливо передавати семантичні залежності у векторних представленнях слів. Перевірити навчену модель можна за допомогою еталонних пар слів з семантичними зв'язками типу «однина-множина», «країна-столиця» та ін. Оцінкою в такому випадку можна вважати відношення правильно визначених пов'язаних слів у парах до загальної кількості перевірочних пар [12].

Мінімальна обробка вхідних даних повинна включати у себе видалення зайвих символів з текстів, наприклад, як переноси рядків. Сам текст повинен розбиватися на токени, наприклад, за допомогою відповідного методу NLTK. Тільки в такому вигляді дані придатні для навчання Word2Vec.

В експериментальній програмі було використано реалізацію Word2Vec у бібліотеці gensim. Оскільки задачею є створення векторних представлень англійських постів із соціальних мереж, то й навчання вкладенням слів потрібно проводити на вибірці постів із соціальних мереж. У таблиці 1 представлені результати обчислення коефіцієнта передачі семантичних зв'язків для різних архітектур моделей у Word2Vec та різних функцій для навчання.

Таблиця 1. Коефіцієнти засвоєння семантичних зв'язків для різних моделей Word2Vec

Алгоритм навчання / архітектура моделі	skip-gram	CBOW
hierarchical softmax	0.04208	0.00203
negative sampling	0.01239	0.00132

З результатів видно, що комбінація архітектури skip-gram з функцією hierarchical softmax дає кращі результати передачі семантичних зв'язків. Слід також зазначити, що для Word2Vec, вибірка в 200 тисяч текстів не є великою, отже, можна досягти кращих результатів навчання цих моделей. Тим не менш, такого результату достатньо для того, щоб створювати векторні представлення постів для подальшої класифікації з точністю (accuracy) 0.6 – 0.6. За 5 експериментів модель skip-gram навчалася в середньому за 21 секунду, а CBOW за 16 секунд.

### Висновки та перспективи подальшого дослідження.

Було проведено ряд експериментів з моделями Word2Vec. З'ясовано, що модель skip-gram видає найвищу точність при великому обсязі навчальних текстів. При цьому, дана модель, при використанні функції hierarchical softmax для навчання на постах з соціальних мереж, також надає найвищу точність у визначенні семантично пов'язаних слів. Проте CBOW є менш обчислювальною витратною, й видає результати подібної, але меншої точності.

На якість навчання моделей Word2Vec може впливати безліч факторів. Обрані методи попередньої обробки даних потенційно впливають на склад навчальних даних та на передані таким чином зв'язки між словами. Параметри при навчанні нейронних мереж також мають великий вплив на точність моделей, наприклад, налаштування випадкового викидання частих слів у текстах може позитивним чином вплинути на організацію векторного простору слів. Усе це може стати об'єктом подальших досліджень.

Беручи до уваги отримані результати, можна створювати більш ефективні підходи до реалізації систем аналізу / моніторингу настроїв у соціальних мережах. Такі системи можна використовувати у соціально-економічних дослідженнях, маркетингу, тощо.

### Список бібліографічного опису

1. Sentiment Analysis and Opinion Mining Techniques [Електронний ресурс] : стаття / Bing Liu – 2012 – Режим доступу: <https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>
2. Sentiment analysis algorithms and applications: A survey [Електронний ресурс] : стаття / Medhata, Wala; Hassan, Ahmed – 2014 – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S2090447914000550>
3. Методи автоматичного аналізу тональності контенту у соціальних мережах для виявлення інформаційно-психологічних впливів. [Електронний ресурс] : стаття / Д. В. Шингалов, С. В. Мелешко, Р. М. Минайленко, В. А. Резніченко // Machinery in agricultural production, industry machine building, automation – 2017 – Режим доступу: <http://dspace.kntu.kr.ua/jspui/bitstream/123456789/6737/1/29.pdf>
4. Distributed Representations of Words and Phrases and their Compositionality [Електронний ресурс] : стаття / Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean – 2013 – Режим доступу: <https://arxiv.org/pdf/1310.4546.pdf>

5. Machine learning methods: An overview [Електронний ресурс] : стаття / Ravil Muhamedyev – 2015 – Режим доступу: [https://www.researchgate.net/publication/320550516\\_Machine\\_learning\\_methods\\_An\\_overview](https://www.researchgate.net/publication/320550516_Machine_learning_methods_An_overview)
6. Logistic Regression Relating Patient Characteristics to Outcomes Efficient Estimation of Word Representations in Vector Space [Електронний ресурс] : стаття / Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean – 2013 – Режим доступу: <https://arxiv.org/pdf/1301.3781.pdf>
7. Feature Selection and Weighting in Sentiment Analysis [Електронний ресурс] : стаття / O'Keefe, Tim; Koprinska, Irena – 2006 – Режим доступу: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.709.1463&rep=rep1&type=pdf>
8. Google Code | Archive – Word2Vec [Електронний ресурс]: [Веб-сайт]. – Tool for computing continuous distributed representations of words – 2013 – Режим доступу: <https://code.google.com/archive/p/Word2Vec/>
9. Word2Vec Tutorial Part I: The Skip-Gram Model [Електронний ресурс] : стаття / Alex Minnaar – 2015 – Режим доступу: <http://alexminnaar.com/2015/04/12/word2vec-tutorial-skipgram.html>
10. Word2Vec Tutorial Part II: The Continuous Bag-of-Words Model [Електронний ресурс] : стаття / Alex Minnaar – 2015 – Режим доступу: <http://alexminnaar.com/2015/05/18/word2vec-tutorial-continuousbow.html>
11. Sentiment Analysis with Word Embedding [Електронний ресурс] : стаття / Deho Oscar Blessed, A. William Agangiba, Felix Larbi Aryeh, Jeffery Antwi Ansah – 2018 – Режим доступу: [https://www.researchgate.net/publication/328523241\\_Sentiment\\_Analysis\\_with\\_Word\\_Embedding](https://www.researchgate.net/publication/328523241_Sentiment_Analysis_with_Word_Embedding)
12. Gensim | Documentation | Word2Vec Model | Evaluation [Електронний ресурс] : [Веб-сайт] – Режим доступу: [https://radimrehurek.com/gensim/auto\\_examples/tutorials/run\\_word2vec.html#evaluating](https://radimrehurek.com/gensim/auto_examples/tutorials/run_word2vec.html#evaluating)
13. Syntactic Dependency-based N-grams as Classification Features [Електронний ресурс] : стаття / Features Grigori Sidorov, Francisco Velasquez, Efstathios Stamatatos, Alexander Gelbukh, Liliana Chanona-Hernández – 2013 – Режим доступу: [https://www.cic.ipn.mx/~sidorov/sn\\_grams\\_MICA2012.pdf](https://www.cic.ipn.mx/~sidorov/sn_grams_MICA2012.pdf)
14. Effectiveness of Hierarchical Softmax in Large Scale Classification Tasks [Електронний ресурс] : стаття / Abdul Arfat Mohammed, Venkatesh Umaashankar – 2015 – Режим доступу: <https://arxiv.org/pdf/1812.05737.pdf>
15. Incremental Skip-gram Model with Negative Sampling [Електронний ресурс] : стаття / Nobuhiro Kaji, Hayato Kobayashi – 2015 – Режим доступу: <https://aclanthology.org/D17-1037.pdf>

#### References

1. Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012;
2. Walaa Medhat, Ahmed Hassan, Hoda Korashy, Sentiment analysis algorithms and applications: A survey, Ain Shams Engineering Journal, Volume 5, Issue 4, 2014, Pages 1093-1113, ISSN 2090-4479;
3. Dmitry Shyngalov, Yelyzaveta Meleshko, Roman Mynaylenko, Vitaliy Reznichenko. Methods of automated sentiment analysis on social networks, Machinery in agricultural production, industry machine building, automation, 2017, Col.30, ISSN 2409-9392;
4. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality. arXiv:1310.4546v1;
5. Muhamedyev, Ravil. (2015). Machine learning methods: An overview. CMNT. 19. 14-29;
6. Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean. Logistic Regression Relating Patient Characteristics to Outcomes Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781v3;
7. Tim O'Keefe, Irena Koprinska. Feature Selection and Weighting in Sentiment Analysis. DOI:10.1.1.709.1463;
11. Oscar Blessed, Deho & Agangiba, A. & Aryeh, Felix & Ansah, Jeffery. (2018). Sentiment Analysis with Word Embedding. 1-4. 10.1109/ICASTECH.2018.8506717.
13. Sidorov G., Velasquez F., Stamatatos E., Gelbukh A., Chanona-Hernández L. (2013) Syntactic Dependency-Based N-grams as Classification Features. In: Batyrshin I., Mendoza M.G. (eds) Advances in Computational Intelligence. MICAI 2012. Lecture Notes in Computer Science, vol 7630. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-37798-3\\_1](https://doi.org/10.1007/978-3-642-37798-3_1)
14. Mohammed, A.A., & Umaashankar, V. (2018). Effectiveness of Hierarchical Softmax in Large Scale Classification Tasks. 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 1090-1094.
15. Kaji, N., & Kobayashi, H. (2017). Incremental Skip-gram Model with Negative Sampling. EMNLP.