

# **Довідковий матеріал\* з виконання практичного завдання №1\*\*** навчальної дисципліни: Інформаційні технології в наукових дослідженнях,

для здобувачів третього (освітньо-наукового) рівня вищої освіти  
спеціальності: 073 Менеджмент, 091 Біологія, 162 Біотехнології та  
біоінженерія\*\*

## **Зауваження\***

Матеріал є керівництвом до дії, а не класичним підручником та відповідно не містить загальноприйнятні посилання, що оформлені як список використаних джерел.

Довідковий матеріал виконано з застосуванням гіперссілок до першоджерел мережі Internet. Використання матеріалу як друкованого джерела чи файлу для читання втрачає більшу частину інформативних посилань, що складає передумови для порушення авторських прав першоджерел.

## **Зауваження\*\***

Довідковий матеріал призначений для більш ґрунтовної теоретичної підготовки та формування стійких навичок з використанням мережі Internet та безпосередньо виконання практичних завдань з застосуванням сучасних програмних засобів оброблення даних.

Довідковий матеріал спрямований на усунення протиріч та двійкового тлумачення визначень та термінології різних предметних областей підготовки здобувачів вищої освіти зі змістом визначень та термінології прийнятим в галузі WEB MINING як розвиток DATA MINING. Для досягнення мети в викладенні теоретичних положень властивих галузі DATA MINING у довідковому матеріалі використано 3 мови демонстрації матеріалу:

Англійська – мова сталої термінології DATA MINING, в випадках відсутньої сталої термінології мовою викладання в галузі.

Українська – мова викладання відповідно до освітніх програм навчального закладу.

Російська – мова на якій викладено довідкові матеріали в першоджерелах, що захищено авторськими правами на час запозичення та мова що обрано як базова на офіційних сайтах хмарних сервісів DATA MINING.

## **Зауваження\*\*\***

Методичне забезпечення з навчальної дисципліни згруповано у три підгрупи в залежності від математичної та мовної підготовки здобувачів за змістом освіти за першим та другим рівнем вищої освіти:

1 підгрупа спеціальностей: 032 Історія та археологія, 033 Філософія, 061 Журналістика, 081 Право, 292 Міжнародні економічні відносини – ґрунтовна мовна підготовка з англійської мови.

2 підгрупа спеціальностей: 073 Менеджмент, 091 Біологія, 162 Біотехнології та біоінженерія – висока мовна та математична підготовка.

3 підгрупа спеціальностей: 051 Економіка, 102 Хімія, 104 Фізика та астрономія, 106 Географія – ґрунтовна фундаментальна математична підготовка.

Зміст навчання за дисципліною однаковий для всіх трьох підгруп, відрізняється виключно методичне забезпечення для найбільш повного наповнення підготовки до практичних занять, а також формування та збереження зацікавленості здобувачів вищої освіти щодо навчального матеріалу з інших галузей знань корисним в сучасних мультидисциплінарних дослідженнях.

## **Введення**

Під час вивчення навчальної дисципліни передбачено прищеплення практичних навичок використання найбільш розповсюджених сучасних програмних засобів STATISTICA, Power BI, Power Pivot та електронних таблиць Excel.

Тематика практичних робіт наведено в таблиці.

### Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	Приклади розрахунків на майбутній момент часу для предметної області (практична робота за індивідуальними завданнями).	2
2	Приклади застосування кількісних методів в моделюванні предметної області досліджень(практична робота за індивідуальними завданнями).	4
3	Пошук джерел інформації за допомогою фільтрів та гіперпосилань у власних дослідженнях (практична робота за індивідуальними завданнями).	4
4	Вибір параметру (критерію) ефективності досліджуваної системи. Розрахунки ефективності систем в досліджуваній області (практична робота за індивідуальними завданнями).	4

### Тематика занять (зміст навчальної дисципліни)

#### Змістовий модуль 1. Місце ІТ технологій в сучасних дослідженнях.

#### ТЕМА 1. ІТ ТЕХНОЛОГІЇ В ПЕРЕДБАЧЕННІ ТА ПРОГНОЗУВАННІ.

Заняття 1. Лекція. Основні поняття. Методи прогнозування.

- Методи прогнозування
- Регресійний аналіз
- Інформаційні продукти OS Windows, Android для побудови моделей прогнозування

Заняття 2. Лекція. Процес підготовки і прийняття рішень. Вихідні дані для прогностичного моделювання ситуацій.

- Процес підготовки і прийняття рішень
- Поняття цільової функції в прийнятті рішень
- Вихідні дані для прогностичного моделювання ситуацій

Заняття 3. **Практичне заняття** (практична робота за індивідуальними завданнями).

#### ТЕМА 2. ОСНОВИ МОДЕЛЮВАННЯ СИТУАЦІЙ ТА ВИРОБЛЕННЯ РІШЕНЬ. ОЦІНКА РИЗИКІВ (ОЦІНКА ПОМИЛКОВИХ РЕЗУЛЬТАТІВ) В ПРИЙНЯТТІ РІШЕНЬ. PESTEL-АНАЛІЗ.

Заняття 1. Лекція. Основні поняття. Класифікація. Методи моделювання характерні для ІТ-систем. Моделювання ситуацій. Оцінка ризиків (оцінка помилкових результатів) в прийнятті рішень.

- Методи моделювання характерні для ІТ-систем
- Витоки помилкових результатів в інформаційних системах. Оцінка ризиків (оцінка помилкових результатів)

Заняття 2. Лекція. Поняття PESTEL-аналізу. Складові PESTEL-аналізу.

- Поняття PESTEL-аналізу
- Складові PESTEL-аналізу та їх характеристика
- Розподіл застосування складових PESTEL-аналізу в різноманітних галузях досліджень

Заняття 3. **Практичне заняття** (практична робота за індивідуальними завданнями).

## **Змістовий модуль 2. ІТ технології пошуку, узагальнення і обробки інформації.**

### ТЕМА 3. WEB MINING ЯК РОЗВИТОК DATA MINING.

Заняття 1. Лекція. Класифікація задач Data Mining.

Заняття 2. Лекція. Витяг Web-контенту в процесі інформаційного пошуку. Поняття фільтрів. Пошук Web-документів з урахуванням гіперпосилань.

- Видобуток даних — Data Mining, задачі Data Mining
- Задача класифікації та регресії
- Задача пошуку асоціативних правил, задача кластеризації
- Практичне застосування Data Mining
- Інтернет-технології

Заняття 3. **Практичне заняття** (практична робота за індивідуальними завданнями).

### ТЕМА 4. ОЦІНКА ЕФЕКТИВНОСТІ СИСТЕМ.

Заняття 1. Лекція. Основні поняття. Критерії ефективності систем (досліджень). Цільова функція.

- Критерії ефективності систем (досліджень)
- Цільова функція
- Задачі теорії операцій
- Поняття оцінки ефективності систем методом мінімаксу та максиміну

Заняття 2. Лекція. Тактика ризику. Завдання теорії ігор для кількісної оцінки ефективності систем в наукових дослідженнях.

- Тактика ризику
- Задачі теорії ігор
- Завдання теорії ігор для кількісної оцінки ефективності систем в наукових дослідженнях

Заняття 3. **Практичне заняття** (практична робота за індивідуальними завданнями).

**Довідковий матеріал з виконання практичного завдання:** Приклади розрахунків на майбутній момент часу для предметної області (практична робота за індивідуальними завданнями).

### Мета:

- формування та розвиток практичних навичок використання електронних таблиць Excel для візуалізації результатів досліджень у вигляді часових (рус: временных рядов),
- розвиток практичних навичок використання гіперпосилань для добору довідкового матеріалу в суміжних галузях досліджень та мультидисциплінарних досліджень,
- прищеплення навичок з використання сучасних програмних сервісів та хмарних технологій як інструментів добутку нових знань.

### Загальні відомості з теорії

За посиланням: [Основаы теории прогнозирования](#) | [Современное прогнозирование \(svetunkov.ru\)](#) розглядаються додатково наступний навчальний матеріал захищений авторським правом:

- [Базовые принципы прогнозирования. Наука](#)
- [Базовые принципы прогнозирования. Практика](#)
- [Системы и модели](#)
- [Измерение и шкалы](#)
- [Виды прогнозируемых процессов](#)
- [Уровни и сроки прогнозирования](#)

О прогнозировании в наше время говорят достаточно часто. Оно у многих на слуху. Так одни люди занимаются прогнозированием котировок валют, другие — экономико-политической ситуации в стране, третьи — продаж пива. Прогнозы всплывают каждый день в том или ином виде в различных средствах массовой информации и дезинформации, при этом многие из них уже на следующий день забываются и не сбываются никогда.

Вообще методы прогнозирования многообразны. Не думайте, что прогнозирование — это статистика. Не подумайте, что прогнозирование — это эконометрика. Не надумайте, что прогнозирование — это математика. Прогнозирование — это уже давно значительно больше и обширнее всего этого. Это уже самостоятельная дисциплина. В каких-то условиях действительно имеет смысл прибегнуть к методам математической статистики, в других — к эконометрике, в третьих — вообще к экспертным методам. Вариантов того, как именно дать прогноз — множество, условий, в которых нужно дать прогноз — мириады. Главное же, что отделяет прогнозирование от остальных дисциплин — это взгляд в будущее.

Именно взглядом в будущее прогнозист и отличается от всех прочих аналитиков и исследователей. Прогнозист не может успокоиться, если модель просто хорошо описала данные, он должен быть уверен в том, что модель даст точный прогноз. Он считает, что главный критерий полезности модели — умение предсказывать. Статистические свойства модели для него важны, но не критичны. Объяснение причин произошедших событий — полезно, но не является самоцелью. Какой смысл размышлять о социально-экономической ситуации, вздыхать о том, что всё плохо или всё хорошо, если эксперт не может точно сказать, что нас ждёт в будущем и что с этим можно сделать?

Кстати, а вы знали, что из всех прогнозов, которые даются в открытую видными экономистами / журналистами / политиками / чиновниками / ..., сбывается в лучшем случае процентов пять? Объясняется это либо тем, что люди, давая такие прогнозы, пытаются произвести впечатление на «народ», либо тем, что эти самые эксперты просто не умеют прогнозировать.

К слову, несколько лет назад журнал «The Economist» провёл эксперимент. Они попросили дать экспертный прогноз по социально-экономической ситуации в Великобритании несколько разных групп людей. В первой были лорды, во второй — топ-менеджеры крупных компаний, в третью входили рядовые аналитики, а в четвёртую — обыкновенные уборщики. Как вы думаете, какая из этих групп дала наиболее точный прогноз? Четвёртая! И произошло это, как говорит Пол Гудвин (Paul Goodwin) из-за того, что люди, обременённые знаниями и опытом, излишне усложнили свой прогноз. А, как показали многие эксперименты (например, M3-Competition) более сложный прогноз не обязательно будет более точным.

В общем, прогнозирование — это очень интересная, но при этом и достаточно сложная дисциплина.

Сразу оговорюсь здесь и далее, когда речь заходит о рядах данных, я в первую очередь буду иметь в виду временные ряды. Пространство лично мне не так интересно, как время.

Не будем вдаваться в длительные разъяснения того, что такое прогнозирование, каковы его субъекты / объекты, каким бывает и прочее. На все эти вопросы можно найти ответы во многих учебниках по прогнозированию (в том числе и в нашем: Том 1, Том 2).

Однако есть один важный элемент касаясь прогнозирования, который не все понимают. Даже некоторые специалисты в этой области не отдают этому отчёт. Прогнозирование нужно не для того, чтобы точно сказать, что именно нас ждёт завтра. Зачастую этого просто невозможно сделать. Нет, например, никакой реальной возможности сказать, какой точно будет температура воздуха завтра в 14.37, или сколько мы продадим валенок на следующей неделе. Что можно сделать, используя методы прогнозирования, — так это снизить неопределённость. Если в ваших руках есть прогноз погоды, то на основе него вы можете решить, что надевать. Если у вас есть прогноз продаж, то вы можете запастись нужным количеством валенок, для того, чтобы заработать больше. Конечно же, прогноз не гарантирует, что завтра в 14.37 будет солнечно, температура будет +24 градус по Цельсию, а продадите вы ровно 53 валенка, но без прогноза, принять правильное решение о том, что надевать по крайней мере достаточно затруднительно (к слову, прогноз из головы человека — это тоже прогноз). Так что по большому счёту **главная задача прогнозирования — это снижение неопределённости.**

За посиланням: [2. Методы прогнозирования \(studfile.net\)](http://studfile.net) розглядаються додатково навчальний матеріал захищений авторським правом, що присвячений змісту різноманітних методів прогнозування:

Единого, универсального, метода прогнозирования не существует. В связи с огромным разнообразием прогнозируемых ситуаций имеется и большое разнообразие методов прогнозирования (свыше 150). На рис.1 представлен один из вариантов классификации методов прогнозирования, основанной на индуктивном и дедуктивном подходах.

Из рис.1 видно, что вся совокупность методов прогнозирования может быть представлена двумя группами – в зависимости от степени их однородности:

- простые методы;
- комплексные методы.

Группа простых методов объединяет однородные по содержанию и используемому инструментарию методы прогнозирования (например, экстраполяция тенденций, морфологический анализ и др.).

Комплексные методы отражают совокупности, комбинации методов, чаще всего реализуемые специальными прогностическими системами (например, методы прогнозного графа, система “Паттерн” и др.).

Кроме того все методы прогнозирования поделены еще на три класса:

- фактографические методы;
- экспертные методы;
- комбинированные методы.

В основу их выделения положен характер информации, на базе которой составляется прогноз.

Фактографические методы базируются на фактическом информационном материале о прошлом и настоящем развитии объекта прогнозирования. Чаще всего применяются при поисковом прогнозировании для эволюционных процессов;



Рис. 1. Классификация методов прогнозирования

Экспертные (интуитивные) методы основаны на использовании знаний специалистов-экспертов об объекте прогнозирования и обобщении их мнений о развитии (поведении) объекта в будущем. Экспертные методы в большей мере соответствуют нормативному прогнозированию скачкообразных процессов;

Комбинированные методы включают методы со смешанной информационной основой, в которых в качестве первичной информации наряду с экспертной используется и фактографическая.

В свою очередь, каждый из перечисленных классов также подразделяется на группы и подгруппы. Так, среди фактографических методов выделяются группы:

- статистических (параметрических) методов;
- опережающих методов.

Группа статистических методов включает методы, основанные на построении и анализе динамических рядов характеристик (параметров) объекта прогнозирования. Среди них наибольшее распространение получили экстраполяция, интерполяция, метод аналогий (модель подобия), параметрический метод и др.

Группа опережающих методов состоит из методов, основанных на использовании свойства научно-технической информации опережать реализацию научно-технических достижений. Среди методов этой группы выделяется публикационный, основанный на анализе и оценке динамики публикаций.

Среди экспертных методов выделяют группы по следующим признакам:

- по количеству привлеченных экспертов;
- по наличию аналитической обработки данных экспертизы

Сущность метода экспертных оценок заключается в проведении экспертами интуитивно-логического анализа проблемы с количественной оценкой суждений и формальной обработкой результатов. Получаемое в результате обработки обобщенное мнение принимается как решение проблемы (в данном случае - прогноз).

Центральным этапом экспертного прогнозирования является проведение опроса экспертов. В зависимости от целей и задач экспертизы, существа и сложности анализируемой проблемы, времени, отведенного на опрос и экспертизу в целом, и допустимой их стоимости, а также от подбора участвующих в ней специалистов, выбирается метод опроса:

- индивидуальный или групповой (коллективный);
- личный (очный) или заочный (путем пересылки анкет);
- устный или письменный;
- открытый или скрытый.

За посыланням: [Прогнозирование - статьи с примерами расчета моделей прогнозов в Excel \(4analytics.ru\)](#) розглядаються навчальний матеріал захищений авторським правом, що присвячено застосуванням можливостям вбудованої мови DAX електронних таблиць Excel:

**Как рассчитать доверительный интервал в Excel. Правило трех сигм + применение на практике.**

Из данной статьи вы узнаете:

- Что такое доверительный интервал?
- Как его самостоятельно рассчитать в Excel? Инструкция с пошаговым описанием и файл с примером!
- В чем суть правила 3-х сигм?
- Как можно применить эти знания на практике?

*В наше время из-за переизбытка информации, связанного с большим ассортиментом товаров, направлений продаж, сотрудников, направлений деятельности и т.д., бывает трудно выделить главное, на что, в первую очередь, стоит обратить внимание и приложить усилия для управления. Определение доверительного интервала и анализ выхода за его границы фактических значений — методика, которая поможет вам выделить ситуации, влияющие на изменение тенденций. Вы сможете развивать позитивные факторы и снизить влияние негативных. Данная технология применяется во многих известных мировых компаниях.*

**Расчет прогноза продаж по методу скользящей средней в Excel**

Из данной статьи вы узнаете:

- Для того, чтобы рассчитать прогноз продаж по методу скользящей средней, необходимо:
- Рассчитать коэффициенты сезонности;
- Выбрать период для расчета среднего значения;
- Рассчитать прогноз, т.е. среднее значение умножить на коэффициент сезонности;
- Учесть дополнительные факторы, которые значительно влияют на продажи;

**3 способа оценки точности прогноза и выбора оптимальной модели?**

Из данной статьи вы узнаете:

- Какие способы оценки прогноза вы можете использовать?
- Как выбрать оптимальную модель, которая поможет вам сделать максимально точный прогноз?
- Как рассчитать показатель "Точность прогноза"?

Прогноз по методу экспоненциального сглаживания с трендом и сезонностью Хольта - Винтерса  
Винтерс развил модель экспоненциального сглаживания с трендом Хольта и добавил в неё сезонность. Преимущество данного метода – это возможность сделать прогноз на длительный период. Но для того чтобы сделать прогноз, например, на 1 год, вам понадобятся данные минимум за 2 полных года, а лучше за 3 - 5 полных лет.

Метод Хольта - Винтерса используется для прогнозирования временных рядов, когда в структуре данных есть сложившийся тренд и сезонность.

Из данной статьи вы узнаете:

**Как оценить точность модели и подобрать оптимальные коэффициенты сглаживания.**

Из данной статьи вы узнаете:

- Что делать если мало данных для прогноза - модель экспоненциального сглаживания...
- Как рассчитать прогноз по методу экспоненциального сглаживания в Excel?
- Как оценить его точность?

Прогноз по методу Прогноз по методу экспоненциального сглаживания Хольта – оптимальный вариант прогноза, когда продажи есть только за несколько периодов (месяцев, дней, недель, кварталов) и еще не понятно - существуют ли тенденции роста или падения.

Хольт развил модель простого экспоненциального сглаживания (прочитать статью о модели прогноза "Простое экспоненциальное сглаживание") и добавил в неё тренд.

Метод Хольта используется для прогнозирования временных рядов, когда есть тенденция к росту или падению значений временного ряда. А также для рядов, когда данные есть не за полный цикл, и сезонность еще не выделить (например, за неполный год для прогноза по месяцам).

**Временные ряды и модели прогнозирования**

В данной статье вы узнаете, о различных характеристиках временных рядов и моделях прогнозирования, которые для них подходят.

Для повышения точности прогнозов продаж важно понимать, что разные модели прогнозирования подходят для разных временных рядов с разными характеристиками. К сожалению, нет одной модели, которая бы подошла для расчета прогноза для всех ситуаций - типов временных рядов.

#### **Bootstrap метод - прогноз нерегулярных, редких продаж**

В этой статье вы узнаете о модели прогнозирования нерегулярных, редких продаж Bootstrapping.

В статье по шагам рассмотрено, как рассчитать прогноз продаж по методу Bootstrapping в Excel.

За посиланням: [Обзор методов прогнозирования \(ivan-shamaev.ru\)](http://ivan-shamaev.ru) розглядаються навчальний матеріал захищений авторським правом, що присвячено огляду методів прогнозування, що властиві до застосування в програмних продуктах.

### **Література до підготовки з виконання практичного завдання**

#### **Основна**

1. Rouaud, Mathieu (2013). Probability, Statistics and Estimation. с. 60.
2. ↑ Chiang, C.L, (2003) Statistical methods of analysis, World Scientific. ISBN 981-238-310-7 - page 274 section 9.7.4 "interpolation vs extrapolation"

#### **Додаткова**

Спільна

3. Методи заповнення пропусків даних у задачах прогнозного моделювання соціально-економічних процесів / П. І. Бідюк, О. М. Терент'єв, Т. І. Просянкін-Жарова // Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту: матеріали міжнародної наукової конференції (ISDMCI-2017). – Херсон, ПП Вишемирський В. – С. 185-187. – Бібліогр.: 2 назви.
4. Побудова предикативної моделі прогнозування злочинності на основі методу подібних траєкторій в системі SAS / Р. Щука, С. Іванов, О. Терент'єв, В. Орешко, Я. Яворський, А. Кузенко // Системний аналіз та інформаційні технології: матеріали 19-ї Міжнародної науково-технічної конференції SAIT 2017, м. Київ, 30 травня – 2 червня 2017 р. – Київ, 2016. – С. 335-336. – Бібліогр.: 6 назв.
5. Bidyuk, P. Dynamic processes forecasting and risk estimation under uncertainty using decision support systems [presentation] / P. Bidyuk, O. Terentiev, T. Prosyankina-Zharova // 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), 29 May-2 June 2017, Kyiv, Ukraine, Igor Sikorsky Kyiv Polytechnic Institute. – Kyiv, 2017. – 27 с.
6. Застосування інструментів SAS BASE для дослідження ефективності методів обробки пропусків у вибірках даних з метою підвищення якості прогнозування показників продовольчої безпеки країни / П. І. Бідюк, О. М. Терент'єв, Т. І. Просянкін-Жарова, В. В. Саваст'янов // System analysis and information technology: 19-th International conference SAIT 2017,



Kyiv, Ukraine, May 22 – 25, 2017 : proceedings. – Kyiv : ESC “IASA” NTUU Igor Sikorsky Kyiv Polytechnic Institute, 2017. – Pp. 253–254.

Додаткова

1 підгрупа спеціальностей

7. Ткач С. М. Управління ризиками інвестиційної діяльності в регіоні: теоретичні основи та прикладні аспекти: монографія / С. М. Ткач. — Львів: ДУ «Інститут регіональних досліджень імені М. І. Долишнього НАН України», 2015. — 234 с.

2 підгрупа спеціальностей

8. Davis G. Management Information Systems: Conceptual Foundations, Structure, and Development. — New York: McGraw-Hill, 1974.
9. Edwards J.S. Expert Systems in Management and Administration — Are they really different from Decision Support Systems? // European Journal of Operational Research, 1992. — Vol. 61. — pp. 114—121.

3 підгрупа спеціальностей

10. Зиновьев А. Ю. Визуализация многомерных данных. — Красноярск: Изд. Красноярского государственного технического университета, 2000. — 180 с.