

Государственное образовательное учреждение
"Приднестровский государственный университет им. Т.Г. Шевченко"

Инженерно-технический институт

Кафедра информационных технологий
и автоматизированного управления производственными процессами

УТВЕРЖДАЮ

Заведующий кафедрой ИТиАУПП

 Ю.А. Столяренко

«28» августа 2020 г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

по дисциплине

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ

Направление подготовки

2.09.04.01 Информатика и вычислительная техника

Профиль подготовки

Информационное и программное обеспечение вычислительных систем

Квалификация (степень)

выпускника: **магистр**

Форма обучения: **очная, заочная**

Год набора: **2020 г.**

Разработал:

к.т.н., доцент кафедры ИТиАУПП,

 /В.С. Попукайло

«28» августа 2020 г.

Тирасполь, 2020

Паспорт фонда оценочных средств по учебной дисциплине

1. В результате изучения дисциплины «Интеллектуальный анализ данных» у обучающегося должны быть сформированы следующие компетенции:

Категория (группа) компетенций	Код и наименование	Код и наименование индикатора достижения универсальной компетенции
<i>Профессиональные компетенции и индикаторы их достижения</i>		
Экспертный анализ эргономических характеристик программных продуктов и/или аппаратных средств. Осуществление технического руководства проектно-изыскательскими работами при проектировании объектов, ввод в действие и освоение проектных мощностей. Осуществление руководства разработкой комплексных проектов на всех стадиях и этапах выполнения работ.	ПК-21 Способен осуществлять экспертный анализ эргономических характеристик программных продуктов и/или аппаратных средств	ИД-1 _{ПК-21} Знать: способы осуществления экспертного анализа эргономических характеристик программных продуктов и или аппаратных средств. ИД-2 _{ПК-21} Уметь: осуществлять экспертный анализ эргономических характеристик программных продуктов и или аппаратных средств. ИД-3 _{ПК-21} Владеть: навыками осуществления экспертного анализа эргономических характеристик программных продуктов и или аппаратных средств.

2. Программа оценивания контролируемой компетенции:

Текущая аттестация	Контролируемые модули, разделы (темы) дисциплины их название	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
РУБЕЖНЫЙ КОНТРОЛЬ	Раздел 1 Раздел 2	ПК-21	Лабораторные работы №1-2
РУБЕЖНАЯ АТТЕСТАЦИЯ	Раздел 3 Раздел 4		Лабораторные работы №3-5
Промежуточная аттестация		Код контролируемой компетенции (или ее части)	Наименование оценочного средства
№1		ПК-21	Зачёт

3. Показатели и критерии оценивания компетенции по этапам формирования, описание шкал оценивания

Этапы оценивания компетенции	Показатели достижения заданного уровня освоения компетенции	Критерии оценивания результатов обучения			
		2	3	4	5
Первый этап	ИД-1 _{ПК-21} Знать: способы осуществления экспертного анализа эргономических характеристик программных продуктов и или аппаратных средств	Не знает	Знает основные способы осуществления анализа эргономических характеристик программных продуктов или аппаратных средств	Знает основные способы, но не может применять знания в полной мере в реальных ситуациях	Знает способы осуществления экспертного анализа эргономических характеристик программных продуктов и или аппаратных средств, может применить их на практике
Второй этап	ИД-2 _{ПК-21} Уметь: осуществлять экспертный анализ эргономических характеристик программных продуктов и или аппаратных средств.	Не умеет	Правильно осуществляет анализ эргономических характеристик программных продуктов и или аппаратных средств на базовом уровне.	Умеет осуществлять экспертный анализ эргономических характеристик программных продуктов и или аппаратных средств	Умеет осуществлять экспертный анализ эргономических характеристик программных продуктов и или аппаратных средств на высоком уровне
Третий этап	ИД-3 _{ПК-21} Владеть: навыками осуществления экспертного анализа эргономических характеристик программных продуктов и или аппаратных средств.	Не владеет	Владеет навыками осуществления анализа эргономических характеристик программных продуктов и или аппаратных средств, но не на	Владеет навыками осуществления анализа эргономических характеристик программных продуктов и или аппаратных средств	Владеет навыками осуществления анализа эргономических характеристик программных продуктов и или аппаратных средств на экспертном уровне

Этапы оценивания компетенции	Показатели достижения заданного уровня освоения компетенции	Критерии оценивания результатов обучения			
		2	3	4	5
			экспертном уровне		

4. Шкала оценивания

Согласно Положению «О порядке организации аттестации в ИТИ ПГУ им. Т.Г. Шевченко, итоговая оценка представляет собой сумму баллов, полученных студентом по итогу освоения дисциплины (модуля):

Оценка в традиционной шкале	Оценка в 100-балльной шкале	Буквенные эквиваленты оценок в шкале ЗЕ (% успешно аттестованных)
5 (отлично)	88–100	А (отлично) – 88-100 баллов
4 (хорошо)	70–87	В (очень хорошо) – 80-87баллов
		С (хорошо) – 70-79 баллов
3 (удовлетворительно)	50–69	Д (удовлетворительно) – 60-69 баллов
		Е (посредственно) – 50-59 баллов
2 (неудовлетворительно)	0–49	Гх – неудовлетворительно, с возможной пересдачей – 21-49 баллов
		Г – неудовлетворительно, с повторным изучением дисциплины – 0-20 баллов

Расшифровка уровня знаний, соответствующего полученным баллам, дается в таблице, указанной ниже

А	“Отлично” - теоретическое содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные программой обучения учебные задания выполнены, качество их выполнения оценено числом баллов, близким к максимальному.
В	“Очень хорошо” - теоретическое содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом в основном сформированы, все предусмотренные программой обучения учебные задания выполнены, качество выполнения большинства из них оценено числом баллов, близким к максимальному.
С	“Хорошо” - теоретическое содержание курса освоено полностью, без пробелов, некоторые практические навыки работы с освоенным материалом сформированы недостаточно, все предусмотренные программой обучения учебные задания выполнены, качество выполнения ни одного из них не оценено минимальным числом баллов, некоторые виды заданий выполнены с ошибками.
Д	“Удовлетворительно” - теоретическое содержание курса освоено частично, но пробелы не носят существенного характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, большинство предусмотренных программой обучения учебных заданий выполнено, некоторые из выполненных заданий, возможно, содержат ошибки.
Е	“Посредственно” - теоретическое содержание курса освоено частично, некоторые практические навыки работы не сформированы, многие предусмотренные программой обучения учебные задания не выполнены, либо качество выполнения некоторых из них оценено числом баллов, близким к минимальному.
ГХ	“Условно неудовлетворительно” - теоретическое содержание курса освоено частично, необходимые практические навыки работы не сформированы, большинство предусмотренных про-

	граммой обучения учебных заданий не выполнено, либо качество их выполнения оценено числом баллов, близким к минимальному; при дополнительной самостоятельной работе над материалом курса возможно повышение качества выполнения учебных заданий.
Ф	“Безусловно неудовлетворительно” - теоретическое содержание курса не освоено, необходимые практические навыки работы не сформированы, все выполненные учебные задания содержат грубые ошибки, дополнительная самостоятельная работа над материалом курса не приведет к какому-либо значимому повышению качества выполнения учебных заданий.

5. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций при изучении учебной дисциплины в процессе освоения образовательной программы

5.1 Образец индивидуального задания к лабораторной работе №1

Проведите исследовательский анализ данных предложенной Базы данных. В качестве отчёта о работе оформите `ipynb notebook` или `Rmd` файл, в котором будут ответы на следующие вопросы:

1. Информация о каком количестве заказов представлена в Базе данных?
2. Назовите продукт, который продавался более всего в натуральном выражении (штуках)?
3. Назовите продукт, выручка которого максимальна?

5.2. Образец индивидуального задания к лабораторной работе №2

1. Постройте `barplots` (столбчатые диаграммы) для визуализации количества заказов и общей выручки по регионам (2 графика).
2. Постройте гистограммы и графики плотности распределения (`density plot`) для количества (`Quantity`) товаров в заказе и для выручки в заказе (не забудьте учесть количество и дисконт при расчёте выручки) (2 графика)
3. Постройте диаграмму рассеяния (`scatter plot`) для отношения количества товаров и выручки в заказе (1 график)
4. Постройте `barplots` для распределения количества заказов по месяцам и по уровням дисконтов. (2 графика)
5. Постройте скрипичные графики (`violin plot`) для распределения выручки по уровням дисконтов (1 график)

5.3 Образец индивидуального задания к лабораторной работе №3

1. Имеет ли размер скидки статистически значимое влияние на количество товара в заказе? Если да, то на каком уровне (-ях) скидки?
2. Как скидка влияет на показатели категории? Влияет ли скидка на выручку по отдельным категориям?
3. Приносит ли каждый регион одинаковый средний доход от заказов?
4. Влияет ли сезонность на доход в какой-либо из категорий?

5.4 Образец индивидуального задания к лабораторной работе №4

1. Какой регуляризатор (`Ridge` или `Lasso`) агрессивнее уменьшает веса при одном и том же α ?
2. Что произойдет с весами `Lasso`, если α сделать очень большим? Поясните, почему так происходит.
3. Какой из регуляризаторов подойдет для отбора неинформативных признаков?

4. Выберите 3 признака с наибольшими по модулю отрицательными коэффициентами (и выпишите их), посмотрите на соответствующие визуализации. Видна ли убывающая линейная зависимость?
5. Выпишите признаки с коэффициентами, близкими к нулю ($< 1e-3$). Как вы думаете, почему модель исключила их из модели?

5.5 Образец индивидуального задания к лабораторной работе №5

1. Используйте в BaggingClassifier параметры по умолчанию, задав только количество деревьев равным 100. Качество классификации новой модели - среднее значение cross_val_score. Сравните работу композиции деревьев с одним решающим деревом.
2. Изучите параметры BaggingClassifier и выберите их такими, чтобы каждый базовый алгоритм обучался не на всех d признаках, а на $\sqrt{d}$ случайных признаков. Каково качество работы алгоритма?
3. Уберите выбор случайного подмножества признаков из BaggingClassifier и добавьте его в DecisionTreeClassifier. Попробуйте выбирать опять же $\sqrt{d}$ признаков. Какое теперь качество полученного классификатора?
4. Изучите, как качество классификации на данном датасете зависит от количества деревьев, количества признаков, выбираемых при построении каждой вершины дерева, а также ограничений на глубину дерева.

5.6 Вопросы к зачёту по дисциплине «Интеллектуальный анализ данных»

1. Законы распределения случайных величин.
2. Идея p -уровня значимости.
3. Корреляционный анализ. Цели и задачи.
4. Критерии согласия.
5. Гетероскедастичность. Мультиколлинеарность.
6. Регрессионный анализ. Цели и задачи.
7. Свойства статистических оценок.
8. Оценка качества линейных моделей.
9. Понятие мощности статистических критериев. Виды ошибок.
10. Критерии однородности.
11. Дисперсионный анализ. Цели и задачи
12. Параметрические и непараметрические статистические критерии.
13. Логистическая регрессия.
14. Задача классификации
15. Задача кластеризации
16. Основные типы визуализации одномерных и двумерных данных
17. Модели, основанные на деревьях решений
18. Ансамбли моделей