Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Новосибирский государственный технический университет»
Кафедра вычислительной техники

Паспорт зачета

по дисциплине «Анализ и интерпретация данных», 3 семестр

1. Методика оценки

Зачет проводится в устной форме, по билетам. Билет формируется по следующему правилу: выбирается два вопроса из диапазона вопросов 1-23 (список вопросов приведен ниже). Таким образом, проверяются результаты обучения по дисциплине, соотнесенные с установленными в программе индикаторами достижения компетенций. В ходе зачета преподаватель вправе задавать студенту дополнительные вопросы из общего перечня (п. 4).

Форма билета для зачета

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
Факультет АВТФ

Билет № _____
к зачету по дисциплине «Анализ и интерпретация данных»

---

1. Вопрос 1
2. Вопрос 2.
3. Задача.

Утверждаю: зав. кафедрой ____ _____ должность, ФИО
(подпись)
(дата)

2. Критерии оценки результатов обучения, соотнесенных с уровнями освоения индикаторов достижения компетенций

Ответ на билет для зачета засчитывается на **продвинутом** уровне, если студент проводит сравнительный комплексный анализ материала, выявляет проблемы, предлагает механизмы их решения, представляет количественные характеристики определенных процессов, приводит конкретные примеры. Совокупность результатов обучения по дисциплине и соотнесенных с ними индикаторов достижения компетенций не содержит пробелов. Установленные в программе компетенции сформированы в полном объеме. Оценка составляет *от 18 до 20 баллов.*

Ответ на билет для зачета засчитывается на **базовом** уровне, если студент при ответе на вопросы формулирует основные понятия, дает характеристику процессов, явлений,

проводит анализ причин, условий, способен представить количественные и качественные характеристики процессов. Совокупность результатов обучения по дисциплине и соотнесенных с ними индикаторов достижения компетенций содержит несущественные пробелы. Установленные в программе компетенции сформированы на базовом уровне. Оценка составляет *от 15 до 17 баллов.*

Ответ на билет для зачета засчитывается на **пороговом** уровне, если студент при ответе на вопросы дает определение основных понятий, может показать причинно-следственные связи явлений. Совокупность результатов обучения по дисциплине и соотнесенных с ними индикаторов достижения компетенций содержит пробелы. Установленные в программе компетенции сформированы на пороговом уровне. Оценка составляет *от 10 до 14 баллов.*

Ответ на билет для зачета считается **неудовлетворительным** (ниже порогового уровня), если студент при ответе на вопросы не дает определений основных понятий, не способен показать причинно-следственные связи явлений. Совокупность результатов обучения по дисциплине и соотнесенных с ними индикаторов достижения компетенций содержит существенные пробелы. Установленные в программе компетенции не сформированы. Оценка составляет *менее 10 баллов.*

3. Шкала оценки

Зачет считается сданным, если сумма баллов по всем заданиям билета составляет от 10 до 20 баллов включительно. Сумма менее 10 баллов признается неудовлетворительным результатом промежуточной аттестации по дисциплине.

В общей оценке по дисциплине баллы за зачет учитываются в соответствии с правилами балльно-рейтинговой системы, установленными в НГТУ.

4. Вопросы к зачету по дисциплине «Анализ и интерпретация данных»

1. Основные понятия АД, типы задач АД, примеры задач. Математический аппарат АД.
2. Основные этапы компьютерной технологии анализа данных, особенности.
3. Графический разведочный анализ данных. Методы визуализации данных. Достоинства и недостатки визуальных методов анализа данных.
4. Первичный разведочный анализ данных (ПРАД). Основные числовые характеристики данных.
5. Первичный разведочный анализ данных (ПРАД). Исследование закона распределения данных. Критерии согласия.
6. Основные методы поиска ассоциативных правил. Характеристики качества правил ассоциации.
7. Кластерный анализ. Постановка задачи. Примеры задач АД. Методологические этапы проведения кластерного анализа. Используемые меры расстояния между объектами и выбор меры. Достоинства и недостатки различных мер расстояний.
8. Кластерный анализ. Классификация методов кластерного анализа данных. Иерархические агломеративные методы (метод одиночной связи, метод полной связи, метод Варда), метод К-средних. Достоинства и недостатки методов кластеризации.
9. Кластерный анализ. Оценка адекватности полученных решений: кофенетическая корреляция; тесты значимости признаков на основе дисперсионного анализа; метод повторных выборок; процедура Монте-Карло.
10. Деревья классификации (деревья решений). Примеры задач АД. Основные понятия, связанные с деревьями классификации, показатели сложности и точности деревьев решений. Методологические этапы построения дерева решений.

11. Деревья классификации (деревья решений). Выбор критерия точности прогноза. Метод CART построения дерева решений.
12. Деревья классификации (деревья решений). Определение момента прекращения ветвления, правила останова. Процедуры проверки прогнозной способности построенного дерева решений.
13. Деревья классификации (деревья решений). Выбор размера дерева, кросс-проверочное отсечение по минимальной цене сложности.
14. Анализ и прогнозирование временных рядов. Примеры задач АД. Определение временного ряда. Классификация факторов, влияющих на формирование значений временного ряда, структурная модель временного ряда.
15. Анализ и прогнозирование временных рядов. Этапы решения задачи прогнозирования временного ряда на примере.
16. Анализ и прогнозирование временных рядов. Методы определения вида функции тренда.
17. Анализ и прогнозирование временных рядов. Гармонический анализ для определения вида сезонной составляющей временного ряда.
18. Анализ и прогнозирование временных рядов. Анализ случайных остатков. Определение адекватности построенной модели (верификация модели) данным наблюдения.
19. Генетические алгоритмы. Определение, задачи АД, решаемые на основе использования генетических алгоритмов. Основные понятия и термины, принятые в области генетических алгоритмов. Характеристики генетических алгоритмов.
20. Эволюционное моделирование. Основные определения. Применение алгоритмов эволюционного моделирования для решения задач структурной идентификации модели.
21. Компьютерные технологии решения разных типов задач АД в современном статистическом программном обеспечении на примере реальных пакетов и программ.
22. Технологии Data Mining, основные этапы, особенности.
23. Направления развития методов, технологий и средств решения задач АД.