

**ВЪПРОСНИК ЗА ДЪРЖАВЕН ИЗПИТ
МАГИСТЪРСКА ПРОГРАМА
МОДЕЛИРАНЕ НА ГОЛЕМИ ДАННИ В БИЗНЕСА И
ФИНАНСИТЕ /BUSINESS ANALYTICS/**

Въпросникът за държавен изпит е утвърден с Протокол № 12/27.06.2022 г. на Факултетния съвет на Стопанския факултет и влиза в сила за випуск 2021/2022 г.

1. Големи данни. Определение. Въведение. Статистически характеристики на многомерни данни. Описание. Пресмятане в средата на Python и R/RStudio. Стъпки за предварителен анализ на данните. Алгоритми за обработка на липсващи наблюдения.
2. Линейна регресия – използване, приложения, графики. Разделяне на данните на тренировъчно и тестово подмножество и оценка на моделите върху тестовото подмножество. Алгоритми за реализация на линейна регресия – синергетичен ефект. Примери в Python и R/Rstudio.
3. Логистична регресия. Класифициране на множеството от наблюдения. Методики за избор на тренировъчно подмножество. Критерии за оценка на моделите – confusion matrix и classification report. Примери в Python и R/Rstudio.
4. Модели за класифициране на данните. Модел на поддържащите вектори (Support Vector Machines). Видове ядра. Сравняване. Параметри на моделите. Алгоритми, реализация, примери в Python.
5. Методики за моделиране на множества от наблюдения чрез модели за класифициране на наблюдения.
6. Намаление на размерността (dimension reduction). Приложение на анализ на главните компоненти. Реализация в Python или R/RStudio.
7. Линеен дискриминантен анализ. Сравнение с модела на главните компоненти. Прилики и разлики. Пакети за реализация в Python или R/RStudio.
8. Моделиране и прогнозиране на едновариантни времеви редове. Автокореелация, стационарност, идентификация на моделите ARMA, сезонност, тестове за Random Walk (слаба пазарна ефективност), фалшива регресия, тестове за единични корени. ARIMA модели. Оценка на параметрите и идентификация на моделите. Приложение на ARMA и ARIMA моделите за прогнозиране на финансови и икономически данни – примери.
9. Моделиране на волатилността. Условна хетероскедастичност във финансовите времеви редове. Стохастична волатилност. ARCH и GARCH модели. Приложения и примери. Алтернативни подходи.

10. Технологична инфраструктура за бизнес разузнаване. Системи за управление на бази от данни, аналитични платформи, външни източници на данни.
11. Процес по разработка на аналитичен отчет. Основни участници, дейности, съпътстващи документи и резултати.
12. Дюрация – същност и приложение на показателя при инвестиции в дългови ценни книжа.
13. Реални опции – характеристики и основни методи за оценка.
14. Предимства и ограничения на NoSQL платформите за управление на бази от данни в сравнение с традиционните релационни бази от данни.
15. Одит на знания. Планиране на одит на знания. Методи за анализ. Критични бизнес процеси.
16. Технологии за управление на знания. Връзка с процесите за управление на знания. Подпомагане на бизнес процесите. Бизнес интелигентност.
17. С какво се различава обучението с утвърждение (reinforcement learning) от останалите клонове на алгоритмичното учене от данни (machine learning)?
18. Опишете алгоритъма SARSA.
19. Направете описание на средата CartPole-v1 от библиотеката Gym. Какви са възможните действия, състояния и награди?

Забележка. Държавният изпит се състои в развиване на два теоретични въпроса и решаване на една практическа задача (на R или Python) по въпросите от въпросника.

Литература:

1. Гурова, Е., А. Антонова, Р. Николов (ред.), *Управление на знания*, Булвест 2000, София, 2012
2. Гурова, Е., Лекции по управление на знания, <http://elearn.uni-sofia.bg/>
3. Иванов, Ив., Вл. Танов, Моделиране на големи данни: методики и алгоритми. *Machine Learning*, Издателство Авангард Прима, София, 2020
4. Evans, J., *Business Analytics*, Pearson; 3rd edition, ISBN-10: 0135231671, 2019
5. Guidolin, M., M. Pedio. *Essentials of Time Series for Financial Applications*. Elsevier Inc. ISBN: 978-0-12-813409-2. 2018.
6. James, G., D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*, Springer, Second Edition 2021. <https://www.statlearning.com/>
7. Jennex, Murray Eugene, *Knowledge Management, Innovation, and Entrepreneurship in a Changing World*, IGI Global, 2020
8. Gym (openai.com) <https://www.gymnasium.ml/>
9. gym/cartpole.py at master · openai/gym · GitHub (https://github.com/openai/gym/blob/master/gym/envs/classic_control/cartpole.py)
10. Lachev, T., *Applied Microsoft Power BI* (4th Edition): Bring your data to life; ISBN-10: 0976635399, 2019
11. Massingham, P., *Knowledge Management: Theory in Practice*, SAGE Publications Ltd; 2019.

12. Mertins, K., P.Heisig, J.Vorbeck, *Knowledge Management – Concepts and Best Practices*, Springer Science & Business Media, 2013
13. Mitchell, T., *Machine Learning*, McGraw-Hill Science/Engineering/Math, 1997
14. Ravichandiran, S. *Deep Reinforcement Learning with Python*, 2020
15. Russo, M., *The Definitive Guide to DAX: Business intelligence for Microsoft Power BI, SQL Server Analysis Services, and Excel* (2nd Edition) (Business Skills), ISBN-10: 1509306978, 2019
16. Shekar, S., *Design Knowledge Management System*, Penman Books, 2021
17. Tsay, Ruey S., *An introduction to analysis of financial data with R*. John Wiley & Sons, inc., publication. ISBN 978-0-470-89081-3. 2013.
18. Vernimmen, P, Q. Pascal, M. Dallochio Maurizio, Y. Fur Le, A. Salvi, *Corporate Finance: Theory and Practice*, 2005, John Wiley & Sons, ISBN-13 978-0-470-09225-5
19. Wei, William W.S. *Multivariate Time Series Analysis and Applications*. John Wiley & Sons, Inc. 2

юни 2022 г.

Катедра Статистика и иконометрия