ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ВОПРОСЫ ДЛЯ ПОДГОТОВКИ К ЭКЗАМЕНУ

дисциплины Анализ и прогнозирование временных рядов методами искусственного интеллекта
для направления 09.04.04 Программная инженерия
уровень образования магистратура
профиль подготовки Искусственный интеллект и инженерия данных
форма обучения очная
кафедра-разработчик Системное программирование

Рабочая программа составлена в соответствии с ФГОС ВО по направлению подготовки 09.04.04 Программная инженерия, утверждённым приказом Минобрнауки от 19.09.2017 № 932

Разработчик программы,
доктор физ-мат. наук, доцент
(ученая степень, ученое звание, должность) _____ (подпись) _____

Зав. кафедрой Системное программирование
д.физ-мат.н., проф.
(ученая степень, ученое звание) _____ (подпись) Л. Б. Соколинский

Челябинск

Вопросы для подготовки к зачету по дисциплине

Введение в дисциплину. Основные понятия: временной ряд, подпоследовательность. Основные задачи интеллектуального анализа временных рядов: поиск аномалий, поиск мотивов, поиск по образцу, восстановление пропусков, прогноз.

Поиск аномалий во временных рядах. Понятие диссонанса. Кусочно-агрегатное сжатие (PAA, Piecewise Aggregate Approximation), Символьно-агрегатное кодирование (SAX, Symbolic Aggregate Approximation). Алгоритм HotSAX. Понятие диапазонного диссонанса. Алгоритмы DRAG и MERLIN.

Поиск подпоследовательностей по образцу. Формальное определение и обзор подходов к решению задачи. Алгоритм UCR-DTW. Мера динамической трансформации времени (DTW, Dynamic Time Warping). Z-нормализация подпоследовательностей. Техника нижних границ: границы LBKim, LBKeogh, LBKeoghEC, каскадное применение нижних границ.

Матричный профиль временного ряда и примитивы анализа данных на его основе. Понятие матричного профиля. Алгоритм SCAMP вычисления матричного профиля. Поиск мотивов. Мера схожести MPdist. Поиск сниппетов (типичных подпоследовательностей). Поиск цепочек (эволюционирующих шаблонов).

Восстановление пропусков и прогноз значений временного ряда. Аналитические алгоритмы HotDeck, kNN, REBOM, TKCM. Модель ARIMA. Нейросетевые методы прогнозирования временных рядов: сеть с долговременной и кратковременной памятью (Long short term memory, LSTM), сеть с управляемыми рекуррентными блоками (Gated Recurrent Units, GRU).

Основная литература

1. Aggarwal C.C. Data Mining: The Textbook. Springer, 2015. 746 p. ISBN 978-3-319-14141-1. Chapter 14. Mining Time Series Data, P. 457-493. https://doi.org/10.1007/978-3-319-14142-8
2. Cryer J.D., Chan K.-S. Time Series Analysis with Applications in R. 2nd Edition. 506 p. ISBN: 978-0-387-75958-6

Дополнительная литература

1. Imani S., Madrid F., Ding W., Crouter S.E., Keogh E.J. Introducing time series snippets: a new primitive for summarizing long time series // Data Min. Knowl. Discov. 2020. Vol. 34, no. 6. P. 1713-1743. https://doi.org/10.1007/s10618-020-00702-y
2. Nakamura T., Imamura M., Mercer R., Keogh E.J. MERLIN: Parameter-Free Discovery of Arbitrary Length Anomalies in Massive Time Series Archives // Proceedings of the 20th IEEE International Conference on Data Mining, ICDM 2020, Sorrento, Italy, November 17-20, 2020. IEEE, 2020. P. 1190-1195. https://doi.org/10.1109/ICDM50108.2020.00147
3. Rakthanmanon T., Campana B.J.L., Mueen A., Batista G.E.A.P.A., Westover M.B., Zhu Q., Zakaria J., Keogh E.J. Addressing Big Data Time Series: Mining Trillions of Time Series Subsequences Under Dynamic Time Warping // ACM Trans. Knowl. Discov. Data. 2013. Vol. 7, no. 3. P. 10:1-10:31. https://doi.org/10.1145/2500489
4. Zhu Y., Gharghabi S., Silva D.F., Dau H.A., Yeh C.-C.M., Senobari N.S., Almaslukh A., Kamgar K., Zimmerman Z., Funning G.J., Mueen A., Keogh E.J. The Swiss army

knife of time series data mining: ten useful things you can do with the matrix profile and ten lines of code // Data Min. Knowl. Discov. 2020. Vol. 34, no. 4. P. 949-979. https://doi.org/10.1007/s10618-019-00668-6

5. Yankov D., Keogh E.J., Rebbapragada U. Disk aware discord discovery: finding unusual time series in terabyte sized datasets // Knowl. Inf. Syst. 2008. Vol. 17, no. 2. P. 241-262. https://doi.org/10.1007/s10115-008-0131-9
6. Yeh C.-C.M., Zhu Y., Ulanova L., Begum N., Dau H.A., Silva D.F., Mueen A., Keogh E.J. Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View That Includes Motifs, Discords and Shapelets // Proceedings of the IEEE 16th International Conference on Data Mining, ICDM 2016, December 12-15, 2016, Barcelona, Spain. IEEE, 2016. P. 1317-1322. https://doi.org/10.1109/ICDM.2016.0179
7. Zhu Y., Imamura M., Nikovski D., Keogh E.J. Matrix Profile VII: Time Series Chains: A New Primitive for Time Series Data Mining // Proceedings of the 2017 IEEE International Conference on Data Mining, ICDM 2017, New Orleans, LA, USA, November 18-21, 2017. IEEE, 2017. P. 695-704. https://doi.org/10.1109/ICDM.2017.79