

Литература:

1. *Алексеев О.Г.* Комплексное применение методов дискретной оптимизации. – М.: Наука, 1987. – 248 с.
2. *Фуругян М.Г.* Некоторые алгоритмы решения минимаксной задачи составления многопроцессорного расписания // Известия Российской академии наук. Теория и системы управления. – 2014. – № 2. – С. 50-56.
3. *Костенко В.А., Смелянский Р.Л., Трекин А.Г.* Синтез структур вычислительных систем реального времени с использованием генетических алгоритмов // Программирование. – 2000. – № 5. – С. 63-72.
4. *Brucker P.* Scheduling Algorithms. – Heidelberg: Springer, 2001. – 365 p.
5. *Гончар Д.Р.* Параллельная реализация мультиоценочного алгоритма составления многопроцессорного расписания без прерываний // Некоторые алгоритмы планирования вычислений и методы многокритериальной оптимизации для многопроцессорных систем. – М.: ВЦ РАН, 2014. – С. 21-31.
6. *Тимошевская Н.Е.* Параллельные методы обхода дерева // Математическое моделирование. – 2004. – Т. 16. № 1. – С. 105-114.
7. *Посыпкин М.А., Сигал И.Х., Галимьянова Н.Н.* Алгоритмы параллельных вычислений для решения некоторых классов задач дискретной оптимизации. – М.: ВЦ РАН, 2005. – 43 с.

Волгина О.А.

Выборочный анализ методов обработки качественной информации в количественном прогнозе

Аннотация: В работе рассматриваются выборочные методы использования качественной информации для решения широкого круга задач долгосрочного прогнозирования и анализа. Описываются свойства качественных характеристик, позволяющие им эффективно дополнять количественные методы. Особое внимание обращается на способы выделения и подготовки данных к машинной обработке.

Ключевые слова: долгосрочное прогнозирование, количественные и качественные параметры, экспертная оценка

При формировании долгосрочного прогноза основная проблема, которую сложно решить, заключается в том, что все доступные данные обычно относятся к прошлому или настоящему, а точной и надежной информации для прогнозирования не хватает. Наблюдается следующая особенность: чем больше горизонт прогноза, тем объем информации о будущем состоянии не только уменьшается, но и изменяется его структура. При изменении интервалов между точками формирования количественные методы теряют свою эффективность. Таким образом, если вся информация будет представлена в количественном виде, прогнозирование не будет иметь никакой ценности. Повышение качества долгосрочного прогноза возможно только за счет комбинации качественных и количественных методов.

Характеристики, на основе которых производятся прогнозы в основном, описываются разными «языками»: как правило, для количественных характеристик мы используем обычные числа, которые называются физическими величинами характеристики; для качественных это слова естественного языка, которые называются лингвистическими ценностями. Лингвистические значения могут быть определены также для количественных характеристик; однако для качественных характеристик физические величины не могут быть определены.

На данный момент самым распространенным методом качественного прогнозирования является экспертные оценки. Экспертное заключение часто необходимо при прогнозировании задач из-за отсутствия соответствующей или доступной информации для использования статистических процедур.

Во многих практических приложениях прогнозированию статистические методы могут оказаться неприменимыми, а экспертная оценка, как метод качественного анализа, может служить единственной основой для прогноза [1].

По сравнению с другими методами прогнозирования, у экспертных систем есть разумные перспективы для распространения, но некоторые препятствия очевидны. Психологические риски создают препятствия для экспертных

систем из-за неразрешимых проблем с верификацией результатов. Поскольку проблемы со сложностью заслуживают внимания, первоначальные попытки создания экспертных систем были простыми, даже если это влекло за собой некоторую потерю предсказательной способности [2].

В статье [3] исследуется использование экспертных систем и искусственного интеллекта (в частности, применение нейронных сетей) к прогнозированию недвижимости. В результате авторы пришли к выводу, что более сложные методы не могут дать лучшего результата простых. В статье описывается использование искусственной нейронной сети самого простого уровня. Нейронная сеть используется как вспомогательный элемент для экспертов.

Для проведения качественного анализа после выбора массива статей (документов) требуется предварительная обработка текстовых данных. Цель состоит в том, чтобы извлечь соответствующую информацию из набора новостей и подготовить ее к машинному обучению. Слова и фразы, которые сигнализируют об изменении цены необходимо верно извлечь.

В [4] предложено разделить текстовую предварительную обработку в три основных этапа: извлечение, выбор и представление функций. Шаг извлечения признаков относится к процессу создания списка признаков, которые являются словами или фразами, извлеченными из документов, в достаточной степени описывающие документы.

Это можно сделать разными способами. Метод семантического анализа «мешок слов» самый популярный [5]. Каждый документ или текст выглядит как неупорядоченный набор слов без сведений о связях между ними. Его можно представить в виде матрицы, каждая строка в которой соответствует отдельному документу или тексту, а каждый столбец – определенному слову. Ячейка на пересечении строки и столбца содержит количество вхождений слова в соответствующий документ, либо иное значение, связанное с «весом» – важностью конкретного термина в конкретном документе. Далее семантически пустые термины удаляются, а методы определения корней слов применяются к каждому слову для обработки различных формы слова как единого признака. Документы часто делятся на две (отрицательные и положительные)

или три (отрицательная, нейтральная и положительная) категории, классы в зависимости от их содержания.

Некоторые исследователи используют метод выявления событий на основе заранее составленных терминов предметной области. В статье [6] прогнозируют фондовые рынки, используя информацию, содержащуюся в статьях, опубликованных в Интернете. В качестве исходных данных используются в основном текстовые статьи, появляющиеся в ведущих и наиболее влиятельных финансовых газетах. Текстовые заявления содержат не только эффект (например, падение акций), но и возможные причины события (например, падение запасов из-за ослабления доллара и, как следствие, ослабления казначейских облигаций). Таким образом, использование текстовой информации повышает качество вводимых данных, из которых выбираются значимые позиции.

Еще одним эффективным методом является использование статистической информации в виде частоты упоминаний. В статье [4] показано, как текст из новостных статей может быть использован для прогнозирования внутридневных движений цен финансовых активов с использованием машин опорных векторов. Авторы используют частотный новостной анализ для создания блока данных, используемых впоследствии как основу прогнозирования. Полученные данные показывают, что отслеживание частоты упоминания информации в пресс-релизах может предоставить дополнительную информацию, которая может быть использована для прогнозирования тенденций цен на акции.

После завершения этапов предварительной обработки текста статьи выравниваются по цене и времени, и готовы в машинной обработке.

Таким образом, методы количественного прогнозирования хорошо показывают результат на задаче краткосрочного прогноза, но для среднесрочного и долгосрочного прогнозирования параметров не дают должного результата. При формировании такого рода прогноза необходимо также проводить анализ качественной информации по теме. Существует несколько методов извлечения информации из текстов и документов, подготавливающие данные для машинной обработки. Выбор метода

во многом зависит от стиля представления информации и прогнозируемой величины.

Работа выполнена в рамках темы: «Фундаментальные исследования по направлению «Модели, методы анализа и синтеза структуры и сценариев развития социально-экономических и технических систем управления, повышения их управляемости и безопасности функционирования в условиях неопределенности, структурных возмущений и чрезвычайных ситуаций» № 0052-2019-0011

Литература:

1. *Rowe G., Wright G.* Expert opinions in forecasting: the role of the Delphi technique / Principles of forecasting. – Springer: Boston, MA, 2001. – P. 125-144.

2. *Armstrong J.S., Yokum J.T.* Potential diffusion of expert systems in forecasting // Technological Forecasting and Social Change. – 2001. – Volume 67. Issue 1. – P. 93-103.

3. *Rossini P.* Using expert systems and artificial intelligence for real estate forecasting / Sixth Annual Pacific-Rim Real Estate Society Conference (24-27 January 2000 Sydney, Australia). – URL: https://www.researchgate.net/publication/255480942_Using_Expert_Systems_and_Artificial_Intelligence_For_Real_Estate_Forecasting (дата обращения 10.10.2021).

4. *Mittermayer M.A.* Forecasting intraday stock price trends with text mining techniques / 37th Annual Hawaii International Conference on System Sciences (5-8 January 2004 Big Island, HI, USA). – URL: <https://ieeexplore.ieee.org/document/1265201> (дата обращения 10.10.2021).

5. *San Kim T., Sohn S.Y.* Machine-learning-based deep semantic analysis approach for forecasting new technology convergence // Technological Forecasting and Social Change. – 2020. – Volume 157. – P. 120095.

6. *B. Wüthrich, D. Permunetilleke, S. Leung, V. Cho, J. Zhang, W. Lam.* Daily prediction of major stock indices from textual www data // Hkie transactions. – 1998. – Volume 5. Issue 3. – P. 151-156.