

АНСАМБЛЬ НЕЙРО-ФАЗЗИ СЕТЕЙ ДЛЯ КЛАСТЕРИЗАЦИИ ДАННЫХ ВЫСОКОЙ РАЗМЕРНОСТИ

Жернова П.Е., Бодянский Е.В.

*Харьковский национальный университет радиоэлектроники,
г. Харьков*

Задача кластеризации многомерных наблюдений, последовательно поступающих на обработку, является важным направлением в рамках Data Stream Mining, а для ее решения разработано достаточно большое число различных методов [1, 2]. Здесь следует отметить кластеризующие нейронные сети Т. Кохонена, наилучшим образом приспособленные для обработки информации в online режиме. При этом априорно предполагается, что количество кластеров, на которое разбивается анализируемый массив данных, известно заранее. Поэтому может быть использован подход, основанный на кластеризующих ансамблях, при этом каждый из членов ансамбля спроектирован на различное число возможных кластеров.

В классе процедур нечеткой кластеризации наиболее популярным является вероятностный алгоритм нечеткой кластеризации, основанный на оптимизации целевой функции. Он получил широкое распространение в Data Mining, однако, теряет свою эффективность в задачах обработки данных высокой размерности из-за возникающего эффекта концентрации норм. Для преодоления этого недостатка было предложено использовать, так называемый, полиномиальный фаззификатор и получена процедура известная как fuzzy C-means with polynomial fuzzifier (PFCM). Также была введена адаптивная online версия PFCM, предназначенная для решения задач Data Stream Mining.

Для решения задач кластеризации данных высокой размерности была предложена модификация FCM со взвешиванием каждого из признаков. Предлагаемая система является ансамблем нейро-фаззи самоорганизующихся карт Т. Кохонена, каждая из которых отличается от других количеством нейронов и значением полиномиального фаззификатора. Настройка каждого из членов ансамбля происходит с помощью модифицированного WTM-правила самообучения, при этом в процессе настройки производится автоматическое взвешивание всех компонент обрабатываемых векторов.

Литература:

1. Жернова П.Е. Ядернв нечітка кластеризація даних на основі ансамблю нейронних мереж / П.Е. Жернова, Є.В. Бодянський // Сучасний стан наукових досліджень та технологій в промисловості. – 2018. – No. 4 (6). – P. 42-49. 2. Шмельов Ю.М. Розробка методу класифікації режимів роботи авіаційного двигуна ТВЗ-117 на основі нейромережових технологій / Ю.М. Шмельов, С.І. Владов, О.Ф. Кришан, С.Д. Гвоздік, Л.І. Чижова // Сучасний стан наукових досліджень та технологій в промисловості. – 2018. – No. 4 (6). – P. 93-102.