**Глава 1 Литературный обзор**

Данная глава последовательно даёт обзор источников информации в области кластерного анализа данных, области больших данных и области баз знаний, в качестве которых рассматриваются литература, электронные документы, компьютерные программы.

Рассматривается взаимосвязь

1.1 Роль больших данных в современных информационных технологиях

1.2 Современное состояние области баз знаний

1.3 Современное состояние области кластерного анализа данных

Рассматривая область кластерного анализа данных, нужно отметить многообразие алгоритмов и метрик. На данный момент не существует общепринятой классификации алгоритмов кластеризации, однако можно выделить несколько основных групп, имеющих схожий принцип работы.

Иерархические алгоритмы

- аггломеративные

- дивизимные

- BIRCH

- CURE

- ROCK

Chameleon

Echidna

Разделяющие

- K-means

- K-medoids

- K-modes

- PAM

- CLARANS

- CLARA

- FCM

Плотностные

Сеточные

Моделируемые