Pfa worksheet 2 solutions

Machine learning

1. a

2. d

3. a

4. a

5. b

6. b

7. a

8. d

9. a

10.a

11. d

12. d

13. The **K**-means clustering algorithm is **sensitive to outliers**, because a mean is easily influenced by **extreme values**. **K**-medoids clustering is a variant of **K**-means that is more robust to noises and **outliers**. ... The group of points in the right form a cluster, while the rightmost point is an **outlier**.

## 14. Advantages of k-means

**Relatively simple to implement.**

**Scales to large data sets.**

**Guarantees convergence.**

**Can warm-start the positions of centroids.**

**Easily adapts to new examples.**

**Generalizes to clusters of different shapes and sizes, such as elliptical clusters.**

**15.** The basic **k**-**means** clustering is based on a non-**deterministic algorithm**. This **means** that running the **algorithm** several times on the same data, could give different results. However, to ensure consistent results, FCS Express performs **k**-**means** clustering using a **deterministic** method.

Sql

1. d

2. d

3. a

4. a

5. b

6. c

7. a

8. c

9. a

10. d

11. b

12. c

13. a

14. b

15. a

**STATISTICS**

1. c

2. c

3. d

4. c

5. c

6. b

7. a

8. a

9. d

10. a

11. c

12. d

13. d

14. a

15. d