

6

Referências bibliográficas

- 1 Lo, S. Y. Eleanor. Library of congress classification and dewey decimal classification: a comparison, United College Journal, Vol. 8 (1970), pp. 125-131.
- 2 VICKERY, B.C. Classificação e indexação nas ciências. Rio de Janeiro: BNG/Brasilart, 1980.
- 3 Phillips, W. Howard. A Primer of book classification. London: Association of Assistant Librarians, 1961.
- 4 HAN, J., e KAMBER, M., 2001, "Cluster Analysis". In: Morgan Kaufmann Publishers (eds.), *Data Mining: Concepts and Techniques*, 1 ed., chapter 8, New York, USA, Academic Press.
- 5 Tan, P., Steinbach, M., and Kumar, V. 2005 Introduction to Data Mining, (First Edition). Addison-Wesley Longman Publishing Co., Inc.
- 6 SEBASTIANI, F. (2002). Machine Learning in Automated Text Categorization. ACM Computing Surveys 34, 1, apud Langie (2003:6).
- 7 Masand, B., G. Linoff, and D. Waltz, "Classifying News Stories Using Memory Based Reasoning", 15th Ann Int ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 59–64, 1992.
- 8 Yang, Y., "An Evaluation of Statistical Approaches to Text Categorization", Journal of Information Retrieval , Vol. 1, No. 1–2, pp. 69–90, 1999.
- 9 Yang, Y. and J. P. Pedersen, "A Comparative Study on Feature Selection in Text Categorization", The Fourteenth International Conference on Machine Learning, pp. 412–420, 1997.
- 10 Joachims, T., "Text Categorization with Support Vector Machines: Learning with Many Relevant Features", European Conference on Machine Learning (ECML), 1998.
- 11 Ng, H. T., W. B. Goh, and K. L. Low, "Feature Selection, Perceptron Learning, and a Usability Case Study for Text Categorization", 20th Ann Int ACM SIGIR Conference on Research and Development in Information Retrieval , pp. 67–73, 1997.

- 12 Wiener, E., J. O. Pedersen, and A. S. Weigend, "A Neural Network Approach to Topic Spotting", Proceedings of the Fourth Annual Symposium on Document Analysis and Information Retrieval , 1995.
- 13 Yang, Y. and X. Liu, "A Re-examination of Text Categorization Methods", Proceedings of SIGIR-99, 22nd ACM International Conference on Research and Development in Information Retrieval , Berkeley, US, 1999.
- 14 McCallum, A. and K. Nigam, "A Comparison of EventModels for Naive Bayes Text Classification", AAAI-98 Workshop on Learning for Text Categorization, 1998.
- 15 Koller, D. and M. Sahami, "Hierarchically Classifying Documents Using Very Few Words", Proceedings of 14th International Conference on Machine Learning, pp.170–178, Nashville, US, 1997.
- 16 S. Chakrabarti, Mining the Web: Discovering Knowledge from Hypertext Data, Morgan Kaufmann, 2002. 3.5, 4, 7.
- 17 A. McCallum and K. Nigam, A comparison of event models for Naive bayes text classification, In Proc. of the AAAI-98 Workshop on learning for text categorization, pp. 41-48. 2.2, 4.
- 18 Vapnik (1995), The Nature of Statistical Learning Theory. Springer, Berlin.
- 19 J. B. MacQueen (1967): "Some Methods for classification and Analysis of Multivariate Observations", Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1:281-297.
- 20 Kaufman, L. and P. Rousseeuw, Finding Groups in Data, Wiley, New York, NY, 1990.
- 21 HAN, J., e KAMBER, M., 2001, "Cluster Analysis". In: Morgan Kaufmann Publishers (eds.), Data Mining: Concepts and Techniques, 1 ed., chapter 8, New York, USA, Academic Press.
- 22 S. Salvador, P. Chan, Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms, in: Proceedings of the Sixteenth IEEE International Conference on Tools with Artificial Intelligence, Boca Raton, FL, USA, IEEE, Piscataway, NJ, USA, 2004, pp. 576–584.
- 23 HRUSCHKA, E. R., e EBECKEN, N. F. F., 2001, "A Genetic algorithm for cluster analysis", Submitted to: IEEE Transactions on Evolutionary Computation, January.
- 24 M. Minsky, A Framework for Representing Knowledge. The Psychology of Computer Vision, P. H. Winston (ed.), McGraw-Hill 1975.

- 25 Barbosa, S. D. J., Breitman, K. K., Furtado, A. L. and Casanova, M. A. (2007) Similarity and Analogy over Application Domains. SBBD.
- 26 B. Coppin. Artificial Intelligence Illuminated. Jones and Bartlett Publishers, 2004.
- 27 Lakoff, George e Johnson, Mark, 1980, Metaphors We Live By, The University of Chicago Press, Chicago.
- 28 Rosch, E., Mervis, C.B., Gray, W., Johnson, D., & Boyes-Braem, P., "Basic objects in Natural Categories", Cognitive Psychology, Vol.8, No.3, (July 1976).

Massa de dados sintéticos A

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<frames>
  <frame id="1" nome="Person_A" nomeClasse="Person" classe="Person">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Bernardo]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[25]]></value>
    </slot>
  </frame>
  <frame id="2" nome="Person_B" nomeClasse="Person" classe="Person">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Frederico]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[28]]></value>
    </slot>
  </frame>
  <frame id="3" nome="Person_C" nomeClasse="Person" classe="Person">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Sérgio]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[31]]></value>
    </slot>
  </frame>
  <frame id="4" nome="Person_D" nomeClasse="Person" classe="Person">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Sônia]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[62]]></value>
    </slot>
  </frame>
</frames>
```

Massa de dados sintéticos B

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<frames>
  <frame id="1" nome="Person_A" nomeClasse="Person" classe="Person">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Bernardo]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[25]]></value>
    </slot>
  </frame>
  <frame id="2" nome="Person_B" nomeClasse="Person" classe="Person">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Frederico]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[28]]></value>
    </slot>
  </frame>
  <frame id="3" nome="Person_C" nomeClasse="Person" classe="Person">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Sérgio]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[31]]></value>
    </slot>
  </frame>
  <frame id="4" nome="Person_D" nomeClasse="Person" classe="Person">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Sônia]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[62]]></value>
    </slot>
  </frame>
  <frame id="5" nome="Employee_A" nomeClasse="Employee" classe="Employee">
    <slot attributeName="name" attributeType="string">
      <value><![CDATA[Bruna]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
      <value><![CDATA[25]]></value>
    </slot>
    <slot attributeName="works" attributeType="string">
      <value><![CDATA[Company A]]></value>
    </slot>
  </frame>
</frames>
```

```

        <slot attributeName="area" attributeType="string">
            <value><![CDATA[engineering]]></value>
        </slot>
        <slot attributeName="salary" attributeType="string">
            <value><![CDATA[1000]]></value>
        </slot>
    </frame>
    <frame id="6" nome="Employee_B" nomeClasse="Employee" classe="Employee">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Ana]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[32]]></value>
        </slot>
        <slot attributeName="works" attributeType="string">
            <value><![CDATA[Company A]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[engineering]]></value>
        </slot>
        <slot attributeName="salary" attributeType="string">
            <value><![CDATA[1500]]></value>
        </slot>
    </frame>
    <frame id="7" nome="Employee_C" nomeClasse="Employee" classe="Employee">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Jorge]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[32]]></value>
        </slot>
        <slot attributeName="works" attributeType="string">
            <value><![CDATA[Company A]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[engineering]]></value>
        </slot>
        <slot attributeName="salary" attributeType="string">
            <value><![CDATA[1000]]></value>
        </slot>
    </frame>
    <frame id="8" nome="Employee_D" nomeClasse="Employee" classe="Employee">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Lui]]></value>
        </slot>
    </frame>

```

```

        <slot attributeName="age" attributeType="string">
            <value><![CDATA[41]]></value>
        </slot>
        <slot attributeName="works" attributeType="string">
            <value><![CDATA[Company B]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>
        </slot>
        <slot attributeName="salary" attributeType="string">
            <value><![CDATA[100]]></value>
        </slot>
    </frame>
    <frame id="9" nome="Employee_E" nomeClasse="Employee" classe="Employee">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Gilda]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[27]]></value>
        </slot>
        <slot attributeName="works" attributeType="string">
            <value><![CDATA[Company C]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>
        </slot>
        <slot attributeName="salary" attributeType="string">
            <value><![CDATA[10000]]></value>
        </slot>
    </frame>
    <frame id="10" nome="Employee_F" nomeClasse="Employee" classe="Employee">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Marcos]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[40]]></value>
        </slot>
        <slot attributeName="works" attributeType="string">
            <value><![CDATA[Company D]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>
        </slot>
        <slot attributeName="salary" attributeType="string">
            <value><![CDATA[1200]]></value>
        </slot>
    </frame>

```

```

        </slot>
    </frame>
    <frame id="11" nome="Student_A" nomeClasse="Student" classe="Student">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Alexia]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[20]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>
        </slot>
        <slot attributeName="level" attributeType="string">
            <value><![CDATA[undergraduate]]></value>
        </slot>
        <slot attributeName="fee" attributeType="string">
            <value><![CDATA[1200]]></value>
        </slot>
    </frame>
    <frame id="12" nome="Student_B" nomeClasse="Student" classe="Student">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Diana]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[20]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>
        </slot>
        <slot attributeName="level" attributeType="string">
            <value><![CDATA[undergraduate]]></value>
        </slot>
        <slot attributeName="fee" attributeType="string">
            <value><![CDATA[1200]]></value>
        </slot>
    </frame>
    <frame id="13" nome="Student_C" nomeClasse="Student" classe="Student">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Lili]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[24]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>

```

```

        </slot>
        <slot attributeName="level" attributeType="string">
            <value><![CDATA[graduate]]></value>
        </slot>
        <slot attributeName="fee" attributeType="string">
            <value><![CDATA[1200]]></value>
        </slot>
    </frame>
    <frame id="14" nome="Student_D" nomeClasse="Student" classe="Student">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Marcelo]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[24]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>
        </slot>
        <slot attributeName="level" attributeType="string">
            <value><![CDATA[graduate]]></value>
        </slot>
        <slot attributeName="fee" attributeType="string">
            <value><![CDATA[1000]]></value>
        </slot>
    </frame>
    <frame id="15" nome="Student_E" nomeClasse="Student" classe="Student">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Alice]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[22]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>
        </slot>
        <slot attributeName="level" attributeType="string">
            <value><![CDATA[undergraduate]]></value>
        </slot>
        <slot attributeName="fee" attributeType="string">
            <value><![CDATA[1000]]></value>
        </slot>
    </frame>
    <frame id="16" nome="Student_F" nomeClasse="Student" classe="Student">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Antonio]]></value>

```

```

        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[25]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[law]]></value>
        </slot>
        <slot attributeName="level" attributeType="string">
            <value><![CDATA[graduate]]></value>
        </slot>
        <slot attributeName="fee" attributeType="string">
            <value><![CDATA[1000]]></value>
        </slot>
    </frame>
    <frame id="17" nome="Student_G" nomeClasse="Student" classe="Student">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Marco]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[24]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[engineering]]></value>
        </slot>
        <slot attributeName="level" attributeType="string">
            <value><![CDATA[graduate]]></value>
        </slot>
        <slot attributeName="fee" attributeType="string">
            <value><![CDATA[1000]]></value>
        </slot>
    </frame>
    <frame id="18" nome="Student_H" nomeClasse="Student" classe="Student">
        <slot attributeName="name" attributeType="string">
            <value><![CDATA[Guilherme]]></value>
        </slot>
        <slot attributeName="age" attributeType="string">
            <value><![CDATA[24]]></value>
        </slot>
        <slot attributeName="area" attributeType="string">
            <value><![CDATA[engineering]]></value>
        </slot>
        <slot attributeName="level" attributeType="string">
            <value><![CDATA[graduate]]></value>
        </slot>
        <slot attributeName="fee" attributeType="string">

```

```

        <value><![CDATA[1000]]></value>
    </slot>
</frame>
<frame id="19" nome="Student_I" nomeClasse="Student" classe="Student">
    <slot attributeName="name" attributeType="string">
        <value><![CDATA[Joao]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
        <value><![CDATA[24]]></value>
    </slot>
    <slot attributeName="area" attributeType="string">
        <value><![CDATA[engineering]]></value>
    </slot>
    <slot attributeName="level" attributeType="string">
        <value><![CDATA[graduate]]></value>
    </slot>
    <slot attributeName="fee" attributeType="string">
        <value><![CDATA[800]]></value>
    </slot>
</frame>
<frame id="20" nome="Student_J" nomeClasse="Student" classe="Student">
    <slot attributeName="name" attributeType="string">
        <value><![CDATA[Ane]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
        <value><![CDATA[22]]></value>
    </slot>
    <slot attributeName="area" attributeType="string">
        <value><![CDATA[engineering]]></value>
    </slot>
    <slot attributeName="level" attributeType="string">
        <value><![CDATA[undergraduate]]></value>
    </slot>
    <slot attributeName="fee" attributeType="string">
        <value><![CDATA[1100]]></value>
    </slot>
</frame>
<frame id="21" nome="Student_L" nomeClasse="Student" classe="Student">
    <slot attributeName="name" attributeType="string">
        <value><![CDATA[Carlos]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
        <value><![CDATA[21]]></value>
    </slot>
    <slot attributeName="area" attributeType="string">

```

```

        <value><![CDATA[engineering]]></value>
    </slot>
    <slot attributeName="level" attributeType="string">
        <value><![CDATA[undergraduate]]></value>
    </slot>
    <slot attributeName="fee" attributeType="string">
        <value><![CDATA[1100]]></value>
    </slot>
</frame>
<frame id="22" nome="Student_M" nomeClasse="Student" classe="Student">
    <slot attributeName="name" attributeType="string">
        <value><![CDATA[Karin]]></value>
    </slot>
    <slot attributeName="age" attributeType="string">
        <value><![CDATA[21]]></value>
    </slot>
    <slot attributeName="area" attributeType="string">
        <value><![CDATA[engineering]]></value>
    </slot>
    <slot attributeName="level" attributeType="string">
        <value><![CDATA[graduate]]></value>
    </slot>
    <slot attributeName="fee" attributeType="string">
        <value><![CDATA[1100]]></value>
    </slot>
</frame>
</frames>

```