BIOINFORMATICS
Vol. 00 no. 00 2005
Pages 1–3

This is a title

Corresponding Author [1,*], Co-Author [2] and Co-Author [2*]
[1]Department of XXXXXXX, Address XXXX etc.
[2]Department of XXXXXXXX, Address XXXX etc.
Received on XXXXX; revised on XXXXX; accepted on XXXXX

Associate Editor: XXXXXXX

ABSTRACT
Motivation: Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text.
Results: Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text
Availability: Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text
Contact: name@bio.com

1 INTRODUCTION

The role of a machine learning framework is to learn the association between the inputs and the outputs in a training set. After this association is learnt, the model is used to predict the output of the unseen inputs. When addressing regression problems, a feature that is usually neglected by most frameworks is the certainty of the prediction. While most frameworks can predict an output for a given input, they do not provide the user with notion of how certain the framework is regarding this prediction. Inclusion of such notion can highly benefit many machine tasks by removing false positives that the method is uncertain about the decisions it has made. To address this problem, we will employ Gaussian Process (GP) framework **?** to predict peptide retention time. We will show, how the property in assigning certainty to the predictions provide a ground for removing more faulty predictions. Moreover, we will argue how this framework can provide us with a model with significantly higher performance compared to previously used frameworks for solving the retention time prediction task.

2 APPROACH

The task of retention time prediction, focuses on determining the retention time of a peptide given its amino acid sequence. To obtain a robust model for solving this task, one needs to address the following problems. Most machine learning frameworks require the inputs to come from a vector space. To map the peptide into a vector space by extracting biologically meaningful features **?**, collecting general statistical entities collected from the sequences **?** or using different kernels such as spectrum kernel **?** or string kernel **?**. Once the feature vectors are calculated, a machine learning framework should be selected for solving the prediction problem. The choice this method can highly affect the quality of the prediction. For example in **?**, the authors have chosen Support Vector Regression (SVR) framework **?** to solve the regression problem.

In this paper, we throughly analyze the retention time prediction problem from the machine learning perspective. First, we will look at different methods that can be used for mapping the peptide sequences into vector spaces. In this process, our aim is to determine the pros and cons of different feature extraction techniques in association with large data analysis. Second, we will look at Gaussian Process (GP) **?** as more sophisticated framework for solving the prediction task and compare its performance with the widely used SVR framework. We will also analyze how varying the size of the training set can affect the performance of both models.

[*]to whom correspondence should be addressed

3 METHODS

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al.*, 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al.*, 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al.*, 2000 might want to know about text text text text

- for bulleted list, use itemize
- for bulleted list, use itemize
- for bulleted list, use itemize

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al.*, 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al.*, 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al.*, 2000 might want to know about text text text text Text Text Text

Table 1. This is table caption

head1	head2	head3	head4
row1	row1	row1	row1
row2	row2	row2	row2
row3	row3	row3	row3
row4	row4	row4	row4

This is a footnote

Fig. 1. Caption, caption.

Fig. 2. Caption, caption.

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text.

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text

4 DISCUSSION

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text.

Table 1 shows that Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text. Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text. Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text.

5 CONCLUSION

(Table 1) Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text.

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text

1. this is item, use enumerate
2. this is item, use enumerate
3. this is item, use enumerate

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text.

Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text Text. Figure 2 shows that the above method Text Text Text Text

ACKNOWLEDGEMENT

Text Text Text Text Text Text Text Text. Bofelli *et al*., 2000 might want to know about text text text text

Funding: Text Text Text Text Text Text Text Text.

REFERENCES

Bofelli,F., Name2, Name3 (2003) Article title, *Journal Name*, **199**, 133-154.

Bag,M., Name2, Name3 (2001) Article title, *Journal Name*, **99**, 33-54.

Yoo,M.S. *et al*. (2003) Oxidative stress regulated genes in nigral dopaminergic neurnol cell: correlation with the known pathology in Parkinson's disease. *Brain Res. Mol. Brain Res.*, **110**(Suppl. 1), 76–84.

Lehmann,E.L. (1986) Chapter title. *Book Title*. Vol. 1, 2nd edn. Springer-Verlag, New York.

Crenshaw, B.,III, and Jones, W.B.,Jr (2003) The future of clinical cancer management: one tumor, one chip. *Bioinformatics*, doi:10.1093/bioinformatics/btn000.

Auhtor,A.B. *et al*. (2000) Chapter title. In Smith, A.C. (ed.), *Book Title*, 2nd edn. Publisher, Location, Vol. 1, pp. ???–???.

Bardet, G. (1920) Sur un syndrome d'obesite infantile avec polydactylie et retinite pigmentaire (contribution a l'etude des formes cliniques de l'obesite hypophysaire). PhD Thesis, name of institution, Paris, France.