|  |
| --- |
|  |

**Development of a SVM classifier of protein secondary structure topology into signal, transmembrane and globular structures**

REPORT IN FULFILLMENT OF THE COURSE 'PROJECT IN MOLECULAR LIFE SCIENCE/KB8024' AT THE STOCKHOLM UNIVERSITY

Revant Gupta | Masters in Molecular Techniques in Life Science | 17.03.2017

SVM classifiers are frequently used in biology and have become very sophisticated in terms of optimisation based on the problem. Sci-kit learn, is a popular python library that provides a broad range of machine learning algorithms, ability to create pipelines, various scoring metrics and other useful options. Compared to the basic steps of choosing a kernel function and optimising the hyperparameters, advanced strategies include combining classifiers (SVM or otherwise) in hierarchical or parallel schemes to improve quality. For SVMs, two stage models that are often used to sort data into signal only and transmembrane only classifiers and then using a consensus mechanism to arrive at a prediction.(3,4) However there have been attempts to use better scoring functions to improve prediction quality(3), as well designing custom kernels.(2) Successful classification of signal vs transmembrane peptides remains a challenge due the commonality of features between the two.(1) Two methods that predict transmembrane and signal topologies using a single models, PHOBIUS and PHILIUS are based on HMMS, dynamic Bayesian learning respectively.(5)

|  |  |  |
| --- | --- | --- |
| **Prediction Method** | **Möller** | **TOPDB** |
| MEMSAT-SVM | 78% | 67% |
| OCTOPUS | 69% | 64% |
| MEMSAT3 | 77% | 66% |
| ENSEMBLE | 61% | 51% |
| PHOBIUS | 67% | 62% |
| HMMTOP | 64% | 57% |
| PRODIV | 46% | 37% |
| SVMTOP | 70% | 42% |
| TMHMM | 60% | 56% |
| PHDhtm | 45% | 49% |

# Table3

Add table 3 from 2.

# Look Great Every Time

* Need a heading? On the Home tab, in the Styles gallery, just click the heading style you want.
* Notice other styles in that gallery as well, such as for a quote, a numbered list, or a bulleted list like this one.
* For best results when selecting text to copy or edit, don’t include space to the left or right of the characters in your selection.

## PICTURE PERFECT

You might like the photo on the cover page as much as we do, but if it’s not ideal for your report, it’s easy to replace it with your own.

Just delete the placeholder picture. Then, on the Insert tab, click Picture to select one from your files.

1. <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-10-159>
2. http://link.springer.com/chapter/10.1007/3-540-45665-1\_1
3. https://www.ncbi.nlm.nih.gov/pubmed/17369623
4. <http://people.eng.unimelb.edu.au/shiltona/publications/giw39.pdf>
5. http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000213