

# Documentation of SVMcon 1.0

---

SVMcon is a software to predict protein residue-residue contacts in 8 Angstrom distance threshold. It is one of the most accurate contact predictors in the 7th and 8th editions of Critical Assessment of Techniques for Protein Structure Prediction (CASP7&8). It is the only free contact map prediction software (including source code) available so far.

---

## Reference

J. Cheng and P. Baldi. Improved Residue Contact Prediction Using Support Vector Machines and a Large Feature Set. BMC Bioinformatics. 8113, 2007.

Author Jianlin Cheng, PhD. School of EECS, University of Central Florida.

Free for academic and scientific use only.

Copyright @ Jianlin Cheng

Please do not re-distribute the software without the permission of the author.

## Contact

Jianlin Cheng at [chengji@missouri.edu](mailto:chengji@missouri.edu)

---

## Third-party software used by SVMcon

SVMcon softwre uses svm\_classify program in SVM-light ([svmlight.joachims.org](http://svmlight.joachims.org)) to classify data points.

## Reference to SVM-light

T. Joachims, 11 in Making large-Scale SVM Learning Practical. Advances in Kernel Methods - Support Vector Learning, B. Schölkopf and C. Burges and A. Smola (ed.), MIT Press, 1999.

I thank the author of SVM-light (Dr. Thorsten Joachims) for the permit to include svm\_classify in the SVMcon package.

Here is the license quoted from the SVM-light website:

SVM-light is free only for non-commercial use. It must not be distributed without prior permission of the author. The author is not responsible for implications from the use of this software.

---

#### Installation

1. install secondary structure and solvent accessibility predictor PSpro (PSpro1.1 or PSpro2.0)

(download it from [http://sysbio.rnet.missouri.edu/multicom\\_toolbox/](http://sysbio.rnet.missouri.edu/multicom_toolbox/). easy to install)

2. unzip svmcon.tar.gz

3. do configuration

```
cd svmcon
```

```
open configure.pl
```

```
set $install_dir to svmcon directory
```

```
set $pspro_dir to PSpro directory (full path)
```

```
save configure.pl
```

```
run .configure.pl to configure.
```

installation is done.

---

#### Test Installation

```
cd test
```

```
..binpredict_map.sh T0288.fasta test.map
```

The output test.map should be the same as T0288.map.final

---

Run the program

```
predict_map.sh fasta_file output_file
```

fasta\_file protein sequence in the simple FASTA format

(line1 name; line 2 a plain protein sequence).

Output file includes the predicted residue contacts in CASP format.

Notice

Since there are a lot of support vectors in the SVM model, it takes from tens of minutes to several hours to predict contacts for a protein, depending on sequence length.