

LIST OF FIGURES

- **Figure 1.1:** Triosephosphateisomerase (TIM) barrel (PDB accession code 8TIM), colored from blue (N-terminus) to red (C-terminus).
- **Figure 1.2** SVM classifications of data. Here data is blue and red color dots which is separated by linear planes.
- **Figure 1.3:** SVM non-linear classifier, in this ξ represents the error which causes due to misclassification of data.
- **Figure 1.4:** Amino_acid_count.pl perl script.
- **Figure 1.5:** Training file generated by amino acid count feature extraction perl script.
- **Figure 1.6:** Perl script for dipeptide count feature extraction.
- **Figure 1.7:** Feature selection dipeptide count training file format for svm_learn.
- **Figure 1.8:** Perl script for PSSM profiles.
- **Figure 1.9:** Feature selection PSSM profiles training file input file format for svm_light.
- **Figure 1.10:** TIMPRED webserver architecture.
- **Figure 1.11:** Roc plot for AAC feature area under curve in the graph. 0.8644.
- **Figure 1.12:** ROC plot for dipeptide composition, area under curve is 0.8166.
- **Figure 1.13:** Roc plot for PSSM profiles. Area under curve is 0.932.
- **Figure 1.14:** TIMPRED home page.
- **Figure 1.15:** Tim barrel webserver next page describing about Tim barrel proteins.
- **Figure 1.16:** Submission page: Protein sequence can be submitted in FASTA format.
- **Figure 1.17:** This snapshot gives the detail of next page after submission. This page generates the JOB Id.
- **Figure 1.18:** The snapshot of result page displaying the SVM score.
- **Figure 1.19:** FAQs related to TIMPRED webserver.
- **Figure 1.20s:** TIMPRED webserver help page for users.

LIST OF TABLES

- **Table 1.1:** TIM-Barrel protein Super families.
- **Table 1.2:** Performance of SVM classifiers for various combination of training features, kernels and values.

LIST OF ABBREVIATIONS

TIM: Triosephosphate Isomerase.

SCOP: Structural Classification of Proteins.

CATH: Class Architecture Topology Homologous.

TIMPRED: Triosephosphate Isomerase Prediction Webserver.

α -Helix: Alpha helix.

β -Sheet :Beta sheet.

SVM: Support Vector Machine.

PDB: Protein Data Bank.

ROC: Receiver Operating Curve.

AUC: Area Under Curve.

Svm-C: Support Vector Machine Classification.

RBF: Radial Basis Function.

k-mer: kernel function.

AAC: Amino Acid Composition.

DPC: Dipeptide Composition.

PSSM: Position Specific Scoring Matrix.