# Additional file 2: Visualization of feature sets

This additional file displays the visualization of feature sets using t-Distributed Stochastic Neighbor Embedding

|  |  |
| --- | --- |
| Perplexity=25 | Perplexity = 50 |
| Figure S1a: Visualization corresponding to protein feature vectors comprised from 1-gram embedding vectors with perplexity equal to 25 | Figure S1b: Visualization corresponding to protein feature vectors comprised from 1-gram embedding vectors with perplexity equal to 50 |
|  |  |
| Figure S2a: Visualization corresponding to protein feature vectors comprised from 2-gram embedding vectors with perplexity equal to 25 | Figure S2b: Visualization corresponding to protein feature vectors comprised from 2-gram embedding vectors with perplexity equal to 50 |
| Figure S3a: Visualization corresponding to protein feature vectors comprised from 3-gram embedding vectors with perplexity equal to 25 | Figure S3b: Visualization corresponding to protein feature vectors comprised from 3-gram embedding vectors with perplexity equal to 50 |
| Figure S4a: Visualization corresponding to protein feature vectors comprised from 4-gram embedding vectors with perplexity equal to 25 | Figure S4b: Visualization corresponding to protein feature vectors comprised from 4-gram embedding vectors with perplexity equal to 50 |
| Figure S5a: Visualization corresponding to protein feature vectors comprised from 5-gram embedding vectors with perplexity equal to 25 | Figure S5b: Visualization corresponding to protein feature vectors comprised from 5-gram embedding vectors with perplexity equal to 50 |
|  Figure S6a: Visualization corresponding to protein feature vectors comprised from 1-gram and 2-gram embedding combined vectors with perplexity equal to 25 | Figure S6b: Visualization corresponding to protein feature vectors comprised from 1-gram and 2-gram embedding combined vectors with perplexity equal to 50 |
| Figure S7a: Visualization corresponding to protein feature vectors comprised from 1-gram and 3-gram embedding combined vectors with perplexity equal to 25 | Figure S7b: Visualization corresponding to protein feature vectors comprised from the combination of 1-gram and 3-gram embedding combined vectors with perplexity equal to 50 |
| Figure S8a: Visualization corresponding to protein feature vectors comprised from 2-gram and 3-gram embedding combined vectors with perplexity equal to 25 | Figure S8b: Visualization corresponding to protein feature vectors comprised from the combination of 2-gram and 3-gram embedding combined vectors with perplexity equal to 50 |