



Protein function prediction as approximate semantic entailment

In the format provided by the authors and unedited

Appendix A Evaluation metrics

We use four different measures to evaluate the performance of our models. Three protein-centric measures F_{\max} , S_{\min} and AUPR and one class-centric AUC.

F_{\max} is a maximum protein-centric F-measure computed over all prediction thresholds. First, we compute average precision and recall using the following formulas:

$$pr_i(t) = \frac{\sum_f I(f \in P_i(t) \wedge f \in T_i)}{\sum_f I(f \in P_i(t))} \quad (\text{A1})$$

$$rc_i(t) = \frac{\sum_f I(f \in P_i(t) \wedge f \in T_i)}{\sum_f I(f \in T_i)} \quad (\text{A2})$$

$$AvgPr(t) = \frac{1}{m(t)} \cdot \sum_{i=1}^{m(t)} pr_i(t) \quad (\text{A3})$$

$$AvgRc(t) = \frac{1}{n} \cdot \sum_{i=1}^n rc_i(t) \quad (\text{A4})$$

where f is a GO class, T_i is a set of true annotations, $P_i(t)$ is a set of predicted annotations for a protein i and threshold t , $m(t)$ is a number of proteins for which we predict at least one class, n is a total number of proteins and I is an indicator function which returns 1 if the condition is true and 0 otherwise. Then, we compute the F_{\max} for prediction thresholds $t \in [0, 1]$ with a step size of 0.01. We count a class as a prediction if its prediction score is greater or equal than t :

$$F_{\max} = \max_t \left\{ \frac{2 \cdot AvgPr(t) \cdot AvgRc(t)}{AvgPr(t) + AvgRc(t)} \right\} \quad (\text{A5})$$

S_{\min} computes the semantic distance between real and predicted annotations based on information content of the classes. The information content $IC(c)$ is computed based on the annotation probability of the class c :

$$IC(c) = -\log(Pr(c|P(c))) \quad (\text{A6})$$

where $P(c)$ is a set of parent classes of the class c . The S_{\min} is computed using the following formulas:

$$S_{\min} = \min_t \sqrt{ru(t)^2 + mi(t)^2} \quad (\text{A7})$$

where $ru(t)$ is the average remaining uncertainty and $mi(t)$ is average misinformation:

$$ru(t) = \frac{1}{n} \sum_{i=1}^n \sum_{c \in T_i - P_i(t)} IC(c) \quad (\text{A8})$$

$$mi(t) = \frac{1}{n} \sum_{i=1}^n \sum_{c \in P_i(t) - T_i} IC(c) \quad (\text{A9})$$

AUPR is the area under the average precision (*AvgPr*) and recall (*AvgRc*) curve.

AUC is a class-centric measure where compute AUC ROC per each class and take the average.

Appendix B ELEmbeddings loss functions

We use the Adam optimizer to minimize the following loss function:

$$L = \frac{1}{N} \sum_{i=1}^N BCELoss(y_{c_i}, y'_{c_i}) + L_{NF1} + L_{NF2} + L_{NF3} + L_{NF4} \quad (\text{B10})$$

ELEmbeddings normalizes TBox axioms in one of the following four normal forms:

NF1: $C \sqsubseteq D$, e.g., *binding* (GO:0005488) SubClassOf: *molecular function* (GO:0003674)

NF2: $C \sqcap D \sqsubseteq E$, e.g., *cutinase activity* (GO:0050525) and *biological regulation* (GO:0065007) SubClassOf: *positive regulation of protein kinase B signaling* (GO:0051897)

NF3: $C \sqsubseteq \exists R.D$, e.g., *positive regulation of arginine biosynthetic process* (GO:1900080) SubClassOf: *positively regulates* (RO:0002213) some *arginine biosynthetic process* (GO:0006526)

NF4: $\exists R.C \sqsubseteq D$, e.g., *part of* (BFO:0000050) some *conjugation* (GO:0000746) SubClassOf: *mammary stem cell proliferation* (GO:0002174)

where C, D, E represent classes and R represents a relation in the ontology. We convert the GO axioms into these four normal forms using a set of conversion rules implemented in the JCel reasoner. As a result, we obtain a set of axioms all of which are in one of these four normal forms, and which are equivalent to the axioms in GO. ELEmbeddings uses these normalized GO axioms as constraints and projects each GO class into an n -ball (represented as a center point in n -dimensional space and a radius) and each relation as a transformation within n -dimensional space. We set the dimension n to 1024 so that it is the same as our protein embeddings. The losses for each of the normal forms are:

$$L_{NF1} = \frac{1}{|NF1|} \sum_{c,d \in NF1} \max(0, \|f_\eta(c) - f_\eta(d)\| + r_\eta(c) - r_\eta(d) - \gamma) \quad (\text{B11})$$

This loss goes to zero when the n -ball for class c is inside the n -ball for class d for all axioms of the first normal form.

$$L_{NF2} = \frac{1}{|NF2|} \sum_{c,d,e \in NF2} \max(0, \|f_\eta(c) - f_\eta(d)\| - r_\eta(c) - r_\eta(d) - \gamma) + \max(0, \|f_\eta(c) - f_\eta(e)\| - r_\eta(c) - \gamma) + \max(0, \|f_\eta(d) - f_\eta(e)\| - r_\eta(c) - \gamma) + \max(0, \min(r_\eta(c), r_\eta(d)) - r_\eta(e) - \gamma) \quad (\text{B12})$$

Method 1	Method 2	MFO	BPO
DeepGO-SE	DeepGOCNN	$5 \cdot 10^{-06}$	$4 \cdot 10^{-86}$
DeepGO-SE	DeepGOZero	$3 \cdot 10^{-12}$	$2 \cdot 10^{-169}$
DeepGO-SE	Tale	$5 \cdot 10^{-07}$	$2 \cdot 10^{-185}$
DeepGO-SE	SPROF-GO	$1 \cdot 10^{-03}$	$7 \cdot 10^{-39}$
DeepGO-SE	MLP (ESM2)	1.00	0.10
DeepGO-SE	MLP	$4 \cdot 10^{-20}$	0.00
DeepGO-SE	Naive	$2 \cdot 10^{-28}$	0.00
DeepGOGAT-SE	DeepGOCNN	0.29	$2 \cdot 10^{-76}$
DeepGOGAT-SE	DeepGOZero	$1 \cdot 10^{-05}$	$8 \cdot 10^{-153}$
DeepGOGAT-SE	Tale	0.18	$1 \cdot 10^{-163}$
DeepGOGAT-SE	SPROF-GO	1.00	$3 \cdot 10^{-28}$
DeepGOGAT-SE	MLP (ESM2)	$6 \cdot 10^{-04}$	$4 \cdot 10^{-04}$
DeepGOGAT-SE	MLP	$3 \cdot 10^{-11}$	0.00
DeepGOGAT-SE	Naive	$8 \cdot 10^{-17}$	0.00

Table D1: Wilcoxon signed rank-test p-values after Bonferroni correction for multiple testing using neXtProt dataset.

This loss minimizes the distance between intersection of n -balls for c and d and n -ball for class e .

$$L_{NF3} = \frac{1}{|NF3|} \sum_{r,c,d \in NF3} \max(0, \|f_\eta(c) - f_\eta(r) - f_\eta(d)\| - r_\eta(c) - r_\eta(d) - \gamma) \quad (\text{B13})$$

Here, we translate the n -ball for class d using relation vector r and minimize the non-overlap between the translated n -ball and the n -ball for class c .

$$L_{NF4} = \frac{1}{|NF4|} \sum_{c,r,d \in NF4} \max(0, \|f_\eta(c) + f_\eta(r) - f_\eta(d)\| + r_\eta(c) - r_\eta(d) - \gamma) \quad (\text{B14})$$

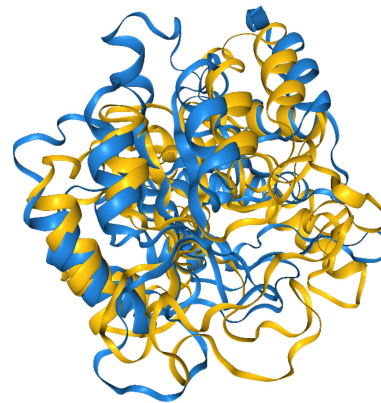
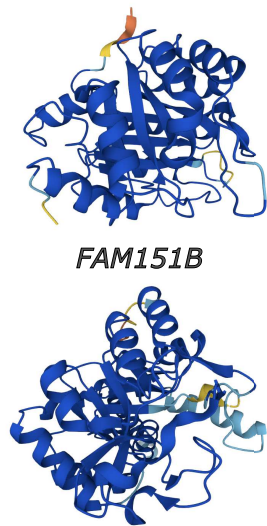
For the fourth normal form, the loss function translates the n -ball for class c using relation vector r and minimizes the distance between the n -balls so that the translated n -ball is inside the n -ball for class d .

We also use a loss for the concept assertion axiom $C(p)$ $L_{C(p)} = \sigma(f_\eta(p) \cdot f_\eta(C)^T + r_\eta(C))$; as the GO does not contain individuals, we use this loss only as in Eqn. ?? for predicting functions of individual proteins.

The parameter γ is a margin parameter and $\|\cdot\|$ is the L_2 norm of the vectors. In addition, we use batch normalization for class embedding vectors to regularize the model and avoid overfitting. The normal form axioms and embeddings are generated jointly for all three sub-ontologies; however, we train and evaluate the function prediction models for each sub-ontology separately.

Appendix C Supplementary figures

Appendix D Supplementary tables



Structural alignment
RMSD: 9.77
TM-Score: 0.42952

Lysophospholipase D GDPD3

Fig. C1: Structural comparison between human proteins *FAM151B* and *Lysophospholipase D GDPD3*. Left: Ribbon diagrams of structures predicted by AlphaFold. The structures are colored according to the AlphaFold confidence score, the predicted local distance difference test (pLDDT), on a scale from 0 to 100. Dark blue: very high (> 90); light blue: high (70–90); yellow: low (50–70); orange: very low (< 50). Right: Structural superimposition by Foldseek, where *FAM151B* is shown in blue and *Lysophospholipase D GDPD3* is shown in gold. Search score is 1.0, RMSD is 9.77 and TM-Score is 0.42952.

Method	F_{\max}			S_{\min}			AUPR			AUC		
	MFO	BPO	CCO	MFO	BPO	CCO	MFO	BPO	CCO	MFO	BPO	CCO
ESM	0.545	0.432	0.721	11.827	39.227	9.358	0.539	0.402	0.724	0.866	0.869	0.918
ESM + GO axioms	0.549	0.426	0.719	11.876	39.749	9.462	0.539	0.394	0.721	0.868	0.859	0.913
ESM + GO PLus axioms	0.552	0.426	0.717	11.750	39.686	9.645	0.550	0.393	0.728	0.867	0.861	0.907
DeepGO-SE	0.554	0.432	0.721	11.681	39.419	9.499	0.552	0.401	0.730	0.874	0.864	0.914
ESM + GAT	0.535	0.432	0.730	11.978	39.201	8.809	0.536	0.404	0.802	0.837	0.878	0.927
ESM + GAT + GO axioms	0.521	0.430	0.727	12.229	39.460	8.743	0.516	0.398	0.733	0.860	0.873	0.927
ESM + GAT + GO PLus axioms	0.517	0.432	0.731	12.321	39.382	8.706	0.513	0.400	0.735	0.855	0.861	0.923
DeepGOGAT-SE	0.525	0.435	0.736	11.137	39.123	8.634	0.523	0.404	0.743	0.861	0.876	0.930
MF + GAT	-	0.453	0.671	-	37.070	9.693	-	0.430	0.721	-	0.833	0.844
MF + GAT + GO axioms	-	0.444	0.666	-	37.737	9.853	-	0.428	0.713	-	0.824	0.831
MF + GAT + GO PLus axioms	-	0.444	0.668	-	37.649	9.803	-	0.426	0.716	-	0.827	0.832
DeepGOGATMF-SE	-	0.448	0.668	-	37.299	9.809	-	0.428	0.679	-	0.831	0.835
MF-Pred + GAT	-	0.455	0.699	-	38.943	9.868	-	0.422	0.760	-	0.864	0.895
MF-Pred + GAT + GO axioms	-	0.441	0.690	-	39.328	10.031	-	0.406	0.696	-	0.852	0.876
MF-Pred + GAT + GO PLus axioms	-	0.443	0.691	-	39.705	10.003	-	0.407	0.749	-	0.853	0.881
DeepGOGATMF-SE-Pred	-	0.444	0.694	-	39.098	9.907	-	0.409	0.753	-	0.855	0.884

Table D2: Ablation study to analyze contributions of GO and GOPlus ontology axioms, PPIs, experimental and predicted MF annotations, and Semantic Entailment to the performance

Table D3: Summary of the UniProtKB/Swiss-Prot dataset

Ontology	GO Terms	Proteins	Groups	Training	Validation	Testing
MFO	6,851	43,279	6,963	52,072	2,964	4,221
BPO	21,356	58,729	9,463	52,584	2,870	3,275
CCO	2,829	59,257	10,019	48,318	4,970	5,969

The table shows the number of GO terms, total number of proteins, number of groups of similar proteins, number of proteins in training, validation and testing sets for the UniProtKB/Swiss-Prot dataset.

Axioms	Strategy	Models	F_{max}			S_{min}			AUPR			AUC		
			MFO	BPO	CCO	MFO	BPO	CCO	MFO	BPO	CCO	MFO	BPO	CCO
GOPlus	min	1	0.526	0.448	0.700	11.071	38.842	9.716	0.516	0.417	0.707	0.864	0.870	0.912
		2	0.530	0.447	0.701	11.029	38.699	9.685	0.522	0.419	0.712	0.867	0.872	0.914
		3	0.530	0.446	0.700	11.010	38.666	9.709	0.523	0.420	0.715	0.868	0.872	0.915
		4	0.529	0.446	0.701	10.997	38.724	9.695	0.524	0.419	0.717	0.869	0.872	0.916
		5	0.529	0.446	0.701	10.999	38.679	9.676	0.524	0.419	0.719	0.874	0.872	0.916
		6	0.531	0.446	0.701	11.007	38.655	9.698	0.524	0.419	0.719	0.874	0.872	0.915
		7	0.529	0.446	0.700	11.047	38.744	9.704	0.524	0.419	0.720	0.874	0.872	0.915
		8	0.527	0.446	0.701	11.046	38.710	9.692	0.524	0.419	0.720	0.875	0.873	0.915
		9	0.528	0.444	0.701	11.044	38.773	9.679	0.524	0.418	0.721	0.874	0.872	0.915
		10	0.527	0.444	0.702	11.053	38.806	9.676	0.523	0.419	0.722	0.874	0.872	0.915
	max	1	0.526	0.448	0.700	11.071	38.842	9.716	0.516	0.417	0.707	0.864	0.870	0.912
		2	0.533	0.448	0.701	10.948	38.678	9.673	0.511	0.419	0.695	0.867	0.871	0.915
		3	0.530	0.448	0.701	10.980	38.731	9.712	0.508	0.421	0.692	0.872	0.870	0.915
		4	0.531	0.449	0.702	11.009	38.743	9.694	0.505	0.421	0.689	0.870	0.869	0.915
		5	0.532	0.448	0.701	10.986	38.755	9.687	0.505	0.419	0.685	0.870	0.869	0.915
		6	0.532	0.448	0.702	10.997	38.818	9.667	0.504	0.419	0.684	0.870	0.870	0.915
		7	0.530	0.448	0.702	11.027	38.810	9.664	0.502	0.419	0.682	0.870	0.870	0.915
		8	0.530	0.448	0.702	11.023	38.782	9.672	0.501	0.419	0.682	0.870	0.870	0.915
		9	0.529	0.448	0.702	11.020	38.782	9.686	0.500	0.419	0.679	0.871	0.870	0.915
		10	0.528	0.448	0.702	11.021	38.774	9.692	0.500	0.419	0.679	0.871	0.870	0.915
	avg	1	0.526	0.448	0.700	11.071	38.842	9.716	0.516	0.417	0.707	0.864	0.870	0.912
		2	0.535	0.448	0.702	10.938	38.605	9.660	0.520	0.421	0.708	0.868	0.872	0.916
		3	0.534	0.450	0.702	10.941	38.592	9.696	0.520	0.423	0.708	0.872	0.872	0.916
		4	0.534	0.449	0.702	10.932	38.613	9.673	0.520	0.423	0.708	0.872	0.872	0.917
		5	0.533	0.450	0.702	10.926	38.613	9.677	0.521	0.422	0.708	0.872	0.872	0.917
		6	0.533	0.449	0.702	10.948	38.579	9.659	0.521	0.421	0.708	0.872	0.873	0.917
		7	0.531	0.450	0.702	10.969	38.609	9.655	0.521	0.422	0.708	0.873	0.873	0.917
		8	0.531	0.449	0.702	10.963	38.595	9.657	0.521	0.422	0.708	0.874	0.873	0.917
		9	0.530	0.449	0.702	10.986	38.602	9.666	0.520	0.422	0.708	0.874	0.873	0.917
		10	0.529	0.449	0.703	10.993	38.606	9.668	0.520	0.422	0.707	0.874	0.873	0.917
GO	min	1	0.523	0.446	0.702	11.137	38.838	9.656	0.508	0.417	0.698	0.871	0.868	0.913
		2	0.519	0.449	0.701	11.190	38.746	9.683	0.514	0.420	0.702	0.876	0.869	0.911
		3	0.518	0.448	0.700	11.186	38.746	9.699	0.511	0.421	0.709	0.875	0.869	0.910
		4	0.520	0.449	0.700	11.176	38.650	9.684	0.515	0.423	0.715	0.875	0.871	0.910
		5	0.519	0.449	0.700	11.153	38.688	9.668	0.515	0.423	0.716	0.876	0.871	0.910
		6	0.520	0.448	0.700	11.150	38.708	9.668	0.516	0.423	0.717	0.875	0.871	0.910
		7	0.520	0.447	0.701	11.135	38.799	9.652	0.517	0.422	0.717	0.876	0.871	0.910
		8	0.520	0.448	0.701	11.141	38.774	9.664	0.518	0.422	0.717	0.877	0.872	0.910
		9	0.519	0.447	0.701	11.170	38.822	9.660	0.517	0.421	0.721	0.877	0.872	0.910
		10	0.519	0.448	0.702	11.170	38.819	9.664	0.518	0.422	0.721	0.876	0.872	0.909
	max	1	0.523	0.446	0.702	11.137	38.838	9.656	0.508	0.417	0.698	0.871	0.868	0.913
		2	0.521	0.447	0.702	11.175	38.803	9.646	0.506	0.418	0.687	0.874	0.869	0.912
		3	0.520	0.448	0.700	11.208	38.773	9.681	0.503	0.420	0.684	0.873	0.869	0.912
		4	0.519	0.451	0.700	11.142	38.728	9.706	0.501	0.422	0.683	0.874	0.870	0.912
		5	0.522	0.451	0.700	11.137	38.754	9.706	0.500	0.422	0.679	0.874	0.870	0.913
		6	0.521	0.450	0.701	11.143	38.707	9.682	0.498	0.423	0.678	0.874	0.870	0.913
		7	0.521	0.450	0.701	11.124	38.726	9.666	0.498	0.423	0.676	0.874	0.869	0.913
		8	0.522	0.450	0.702	11.120	38.729	9.673	0.498	0.423	0.674	0.874	0.870	0.914
		9	0.522	0.450	0.702	11.155	38.744	9.673	0.497	0.423	0.675	0.873	0.870	0.914
		10	0.521	0.450	0.702	11.169	38.717	9.661	0.516	0.422	0.674	0.873	0.870	0.913
	avg	1	0.523	0.446	0.702	11.137	38.838	9.656	0.508	0.417	0.698	0.871	0.868	0.913
		2	0.520	0.450	0.702	11.141	38.741	9.664	0.514	0.420	0.697	0.876	0.870	0.913
		3	0.521	0.450	0.701	11.151	38.653	9.676	0.512	0.422	0.700	0.876	0.870	0.913
		4	0.521	0.453	0.702	11.088	38.574	9.670	0.514	0.425	0.704	0.877	0.872	0.913
		5	0.522	0.453	0.702	11.086	38.625	9.664	0.515	0.425	0.704	0.877	0.872	0.914
		6	0.523	0.453	0.702	11.081	38.577	9.646	0.515	0.425	0.703	0.877	0.872	0.914
		7	0.522	0.452	0.702	11.066	38.591	9.626	0.516	0.425	0.703	0.877	0.872	0.915
		8	0.523	0.452	0.702	11.066	38.585	9.637	0.516	0.425	0.703	0.877	0.872	0.915
		9	0.523	0.452	0.702	11.066	38.590	9.638	0.517	0.425	0.704	0.877	0.872	0.915
		10	0.524	0.451	0.702	11.072	38.586	9.634	0.517	0.425	0.705	0.877	0.872	0.914

Table D4: Performance of semantic entailment depending on axioms, aggregation strategy and number of models for DeepGO-SE model

Axioms	Strategy	Models	F_{max}			S_{min}			AUPR			AUC		
			MFO	BPO	CCO	MFO	BPO	CCO	MFO	BPO	CCO	MFO	BPO	CCO
GO	min	1	0.493	0.448	0.712	11.601	38.115	8.755	0.480	0.423	0.713	0.869	0.876	0.924
		2	0.494	0.448	0.715	11.585	38.076	8.713	0.484	0.424	0.723	0.870	0.877	0.925
		3	0.496	0.449	0.714	11.584	38.079	8.738	0.484	0.425	0.725	0.870	0.877	0.925
		4	0.497	0.449	0.714	11.559	38.096	8.710	0.484	0.425	0.727	0.869	0.878	0.923
		5	0.499	0.449	0.713	11.531	38.077	8.720	0.485	0.425	0.728	0.869	0.878	0.923
		6	0.498	0.448	0.713	11.530	38.109	8.735	0.485	0.424	0.730	0.868	0.878	0.923
		7	0.498	0.448	0.713	11.518	38.117	8.747	0.486	0.423	0.731	0.868	0.878	0.922
		8	0.498	0.448	0.713	11.515	38.123	8.779	0.486	0.424	0.732	0.868	0.878	0.922
		9	0.498	0.448	0.656	11.509	38.158	11.144	0.486	0.424	0.696	0.868	0.878	0.893
		10	0.499	0.448	0.652	11.499	38.163	11.456	0.487	0.424	0.682	0.867	0.878	0.879
	max	1	0.493	0.448	0.712	11.601	38.115	8.755	0.480	0.423	0.713	0.869	0.876	0.924
		2	0.500	0.450	0.715	11.506	38.072	8.706	0.483	0.424	0.706	0.870	0.877	0.926
		3	0.499	0.449	0.714	11.528	38.074	8.726	0.482	0.423	0.703	0.871	0.878	0.928
		4	0.500	0.451	0.715	11.504	38.070	8.718	0.482	0.424	0.699	0.871	0.878	0.928
		5	0.499	0.450	0.714	11.520	38.104	8.727	0.480	0.423	0.697	0.869	0.878	0.927
		6	0.500	0.450	0.714	11.528	38.143	8.731	0.480	0.423	0.696	0.869	0.878	0.927
		7	0.501	0.450	0.713	11.523	38.146	8.785	0.480	0.423	0.693	0.870	0.878	0.928
		8	0.501	0.450	0.713	11.516	38.195	8.792	0.480	0.423	0.691	0.870	0.879	0.928
		9	0.501	0.450	0.694	11.526	38.215	9.095	0.478	0.422	0.667	0.869	0.879	0.898
		10	0.500	0.450	0.694	11.540	38.228	9.115	0.478	0.422	0.664	0.870	0.879	0.894
	avg	1	0.493	0.448	0.712	11.601	38.115	8.755	0.480	0.423	0.713	0.869	0.876	0.924
		2	0.498	0.449	0.715	11.532	38.019	8.684	0.486	0.425	0.718	0.871	0.878	0.927
		3	0.498	0.450	0.715	11.483	37.984	8.681	0.487	0.426	0.719	0.872	0.879	0.928
		4	0.500	0.451	0.715	11.469	37.930	8.664	0.488	0.426	0.720	0.872	0.879	0.928
		5	0.500	0.450	0.715	11.448	37.934	8.672	0.488	0.426	0.720	0.871	0.879	0.928
		6	0.500	0.450	0.715	11.446	37.998	8.655	0.488	0.425	0.721	0.871	0.879	0.928
		7	0.502	0.450	0.715	11.431	37.987	8.678	0.489	0.425	0.720	0.871	0.879	0.929
		8	0.502	0.450	0.715	11.432	37.981	8.691	0.489	0.425	0.720	0.871	0.880	0.929
		9	0.502	0.450	0.714	11.421	37.975	8.710	0.489	0.425	0.724	0.871	0.880	0.918
		10	0.502	0.451	0.712	11.419	37.970	8.731	0.490	0.425	0.769	0.872	0.880	0.915
GOPlus	min	1	0.496	0.454	0.714	11.645	37.990	8.668	0.488	0.430	0.712	0.861	0.875	0.927
		2	0.498	0.453	0.715	11.529	37.973	8.672	0.488	0.429	0.725	0.867	0.877	0.927
		3	0.500	0.451	0.714	11.506	37.976	8.680	0.490	0.427	0.730	0.868	0.877	0.927
		4	0.500	0.452	0.714	11.509	37.948	8.673	0.491	0.429	0.731	0.868	0.878	0.928
		5	0.500	0.452	0.715	11.478	37.959	8.678	0.491	0.429	0.733	0.868	0.877	0.928
		6	0.501	0.452	0.715	11.450	37.996	8.705	0.491	0.429	0.733	0.867	0.877	0.928
		7	0.501	0.452	0.715	11.460	38.005	8.706	0.491	0.429	0.733	0.867	0.877	0.928
		8	0.501	0.452	0.714	11.462	38.024	8.707	0.491	0.428	0.734	0.867	0.877	0.928
		9	0.500	0.451	0.715	11.468	38.026	8.712	0.490	0.428	0.734	0.867	0.877	0.928
		10	0.500	0.451	0.715	11.468	38.041	8.710	0.490	0.428	0.735	0.867	0.878	0.928
	max	1	0.496	0.454	0.714	11.645	37.990	8.668	0.488	0.430	0.712	0.861	0.875	0.927
		2	0.501	0.454	0.714	11.527	37.988	8.672	0.489	0.429	0.708	0.865	0.878	0.928
		3	0.502	0.454	0.715	11.498	37.982	8.665	0.489	0.428	0.703	0.866	0.879	0.928
		4	0.503	0.456	0.715	11.477	37.930	8.662	0.489	0.429	0.701	0.866	0.880	0.929
		5	0.504	0.456	0.715	11.494	37.959	8.683	0.487	0.429	0.699	0.867	0.880	0.929
		6	0.502	0.456	0.715	11.504	37.954	8.692	0.485	0.429	0.697	0.867	0.880	0.929
		7	0.502	0.456	0.715	11.523	37.967	8.726	0.484	0.429	0.696	0.868	0.880	0.929
		8	0.503	0.456	0.715	11.514	37.988	8.732	0.485	0.429	0.695	0.868	0.879	0.928
		9	0.504	0.456	0.715	11.512	37.975	8.741	0.485	0.429	0.694	0.867	0.880	0.928
		10	0.504	0.456	0.715	11.519	37.958	8.742	0.484	0.429	0.693	0.867	0.880	0.928
	avg	1	0.496	0.454	0.714	11.645	37.990	8.668	0.488	0.430	0.712	0.861	0.875	0.927
		2	0.501	0.454	0.715	11.491	37.950	8.609	0.492	0.430	0.719	0.867	0.879	0.928
		3	0.501	0.454	0.716	11.452	37.911	8.629	0.492	0.429	0.720	0.868	0.880	0.929
		4	0.502	0.455	0.716	11.411	37.833	8.616	0.494	0.432	0.722	0.869	0.881	0.929
		5	0.502	0.456	0.717	11.409	37.811	8.607	0.494	0.432	0.722	0.869	0.881	0.929
		6	0.503	0.456	0.717	11.404	37.817	8.618	0.494	0.432	0.722	0.870	0.881	0.929
		7	0.502	0.455	0.717	11.443	37.839	8.624	0.494	0.432	0.722	0.870	0.881	0.929
		8	0.502	0.455	0.717	11.450	37.829	8.643	0.494	0.431	0.722	0.870	0.880	0.929
		9	0.502	0.456	0.717	11.461	37.820	8.639	0.493	0.431	0.722	0.869	0.880	0.929
		10	0.502	0.455	0.716	11.447	37.827	8.654	0.493	0.431	0.722	0.869	0.881	0.929

Table D5: Performance of semantic entailment depending on axioms, aggregation strategy and number of models for DeepGOGAT-SE model

Axioms	Strategy	Models	F_{max}		S_{min}		AUPR		AUC		
			BPO	CCO	BPO	CCO	BPO	CCO	BPO	CCO	
GO	min	1	0.429	0.658	37.441	9.778	0.422	0.661	0.820	0.842	
		2	0.433	0.659	37.336	9.736	0.421	0.667	0.819	0.843	
		3	0.434	0.659	37.351	9.766	0.423	0.668	0.813	0.843	
		4	0.434	0.660	37.360	9.784	0.424	0.669	0.813	0.843	
		5	0.433	0.659	37.463	9.808	0.423	0.669	0.811	0.842	
		6	0.434	0.659	37.470	9.808	0.423	0.670	0.811	0.840	
		7	0.434	0.659	37.570	9.821	0.423	0.670	0.811	0.840	
		8	0.434	0.659	37.596	9.801	0.423	0.671	0.810	0.841	
		9	0.434	0.659	37.602	9.802	0.423	0.671	0.809	0.839	
		10	0.434	0.659	37.614	9.807	0.423	0.672	0.808	0.839	
	max	1	0.429	0.658	37.441	9.778	0.422	0.661	0.820	0.842	
		2	0.436	0.658	37.445	9.749	0.425	0.658	0.824	0.845	
		3	0.443	0.657	37.384	9.839	0.427	0.654	0.825	0.847	
		4	0.445	0.657	37.399	9.833	0.426	0.652	0.825	0.849	
		5	0.446	0.657	37.389	9.818	0.425	0.651	0.824	0.848	
		6	0.447	0.657	37.521	9.824	0.425	0.651	0.825	0.848	
		7	0.447	0.657	37.574	9.843	0.425	0.698	0.824	0.848	
		8	0.448	0.657	37.615	9.851	0.424	0.697	0.824	0.848	
		9	0.448	0.658	37.682	9.848	0.424	0.697	0.823	0.848	
		10	0.448	0.658	37.733	9.857	0.424	0.697	0.823	0.849	
	avg	1	0.429	0.658	37.441	9.778	0.422	0.661	0.820	0.842	
		2	0.431	0.660	37.268	9.726	0.419	0.665	0.824	0.845	
		3	0.433	0.659	37.146	9.743	0.423	0.664	0.825	0.847	
		4	0.432	0.660	37.175	9.752	0.422	0.665	0.825	0.848	
		5	0.432	0.660	37.201	9.753	0.421	0.665	0.825	0.848	
		6	0.432	0.660	37.185	9.765	0.421	0.664	0.826	0.847	
		7	0.432	0.660	37.170	9.768	0.421	0.664	0.826	0.847	
		8	0.432	0.660	37.182	9.765	0.421	0.664	0.826	0.849	
		9	0.432	0.659	37.192	9.771	0.421	0.664	0.825	0.847	
		10	0.432	0.660	37.196	9.763	0.421	0.664	0.824	0.848	
	GOPlus	min	1	0.428	0.658	37.313	9.809	0.414	0.659	0.823	0.847
			2	0.429	0.659	37.344	9.779	0.417	0.665	0.822	0.847
			3	0.431	0.659	37.340	9.789	0.420	0.668	0.821	0.842
			4	0.432	0.659	37.368	9.776	0.420	0.670	0.820	0.841
			5	0.432	0.659	37.440	9.763	0.420	0.671	0.819	0.840
			6	0.432	0.659	37.511	9.755	0.421	0.672	0.818	0.839
			7	0.433	0.660	37.495	9.752	0.421	0.672	0.816	0.839
			8	0.434	0.660	37.511	9.765	0.421	0.673	0.816	0.839
			9	0.434	0.660	37.545	9.773	0.422	0.673	0.815	0.839
			10	0.434	0.660	37.553	9.779	0.422	0.673	0.815	0.839
max		1	0.428	0.658	37.313	9.809	0.414	0.659	0.823	0.847	
		2	0.430	0.657	37.313	9.819	0.413	0.655	0.827	0.851	
		3	0.436	0.657	37.336	9.844	0.413	0.653	0.827	0.852	
		4	0.437	0.657	37.382	9.851	0.411	0.652	0.828	0.852	
		5	0.439	0.657	37.355	9.854	0.410	0.650	0.827	0.851	
		6	0.442	0.657	37.386	9.856	0.412	0.649	0.828	0.852	
		7	0.442	0.658	37.424	9.860	0.412	0.699	0.827	0.852	
		8	0.444	0.657	37.429	9.867	0.411	0.699	0.827	0.851	
		9	0.445	0.657	37.456	9.873	0.411	0.699	0.827	0.850	
		10	0.445	0.657	37.515	9.877	0.411	0.699	0.827	0.850	
avg		1	0.428	0.658	37.313	9.809	0.414	0.659	0.823	0.847	
		2	0.429	0.658	37.263	9.784	0.416	0.662	0.826	0.851	
		3	0.430	0.659	37.205	9.768	0.418	0.662	0.827	0.851	
		4	0.430	0.660	37.279	9.752	0.418	0.664	0.828	0.850	
		5	0.429	0.660	37.254	9.737	0.418	0.663	0.827	0.851	
		6	0.430	0.660	37.227	9.727	0.418	0.664	0.827	0.850	
		7	0.431	0.660	37.182	9.724	0.419	0.663	0.827	0.850	
		8	0.431	0.659	37.153	9.729	0.419	0.664	0.827	0.849	
		9	0.432	0.660	37.164	9.734	0.419	0.663	0.827	0.848	
		10	0.432	0.660	37.174	9.731	0.419	0.663	0.826	0.848	

Table D6: Performance of semantic entailment depending on axioms, aggregation strategy and number of models for DeepGOGATMF-SE model

Axioms	Strategy	Models	F_{max}		S_{min}		AUPR		AUC	
			BPO	CCO	BPO	CCO	BPO	CCO	BPO	CCO
GO	min	1	0.452	0.673	38.317	10.062	0.427	0.685	0.868	0.892
		2	0.451	0.671	38.199	10.087	0.427	0.687	0.869	0.893
		3	0.454	0.671	38.136	10.052	0.429	0.688	0.868	0.891
		4	0.454	0.671	38.086	10.051	0.430	0.689	0.869	0.891
		5	0.453	0.671	38.163	10.062	0.428	0.690	0.869	0.889
		6	0.452	0.671	38.109	10.046	0.428	0.690	0.868	0.888
		7	0.452	0.671	38.123	10.041	0.428	0.691	0.868	0.888
		8	0.453	0.671	38.138	10.046	0.428	0.691	0.868	0.888
		9	0.452	0.671	38.180	10.059	0.428	0.691	0.868	0.888
		10	0.452	0.670	38.185	10.077	0.428	0.690	0.868	0.888
	max	1	0.452	0.673	38.317	10.062	0.427	0.685	0.868	0.892
		2	0.451	0.674	38.405	10.011	0.426	0.681	0.869	0.897
		3	0.454	0.674	38.353	10.036	0.429	0.679	0.870	0.898
		4	0.453	0.673	38.439	10.076	0.429	0.676	0.870	0.900
		5	0.453	0.675	38.587	10.093	0.428	0.676	0.870	0.900
		6	0.453	0.675	38.551	10.127	0.427	0.676	0.870	0.900
		7	0.452	0.674	38.653	10.136	0.426	0.675	0.869	0.900
		8	0.452	0.673	38.742	10.172	0.426	0.674	0.869	0.898
		9	0.452	0.674	38.813	10.164	0.425	0.674	0.869	0.898
		10	0.452	0.673	38.890	10.163	0.424	0.673	0.868	0.898
	avg	1	0.452	0.673	38.317	10.062	0.427	0.685	0.868	0.892
		2	0.452	0.673	38.271	10.007	0.427	0.686	0.870	0.897
		3	0.454	0.674	38.188	9.990	0.430	0.686	0.871	0.898
		4	0.454	0.673	38.177	9.987	0.430	0.686	0.871	0.899
		5	0.454	0.673	38.201	9.992	0.430	0.687	0.872	0.900
		6	0.454	0.674	38.124	10.007	0.429	0.686	0.872	0.900
		7	0.454	0.674	38.092	9.998	0.429	0.686	0.872	0.899
		8	0.454	0.674	38.121	9.996	0.429	0.686	0.872	0.899
		9	0.454	0.673	38.171	9.991	0.429	0.686	0.872	0.899
		10	0.454	0.673	38.187	9.996	0.429	0.686	0.871	0.898
GOPlus	min	1	0.452	0.669	38.383	10.050	0.428	0.729	0.871	0.897
		2	0.452	0.673	38.380	9.960	0.430	0.690	0.872	0.896
		3	0.453	0.673	38.305	9.970	0.431	0.691	0.872	0.896
		4	0.453	0.674	38.262	9.963	0.431	0.692	0.873	0.896
		5	0.454	0.673	38.224	9.957	0.431	0.692	0.873	0.896
		6	0.454	0.673	38.210	9.955	0.431	0.692	0.872	0.897
		7	0.454	0.673	38.203	9.951	0.431	0.693	0.872	0.897
		8	0.454	0.673	38.222	9.960	0.431	0.693	0.871	0.897
		9	0.454	0.673	38.182	9.978	0.431	0.693	0.872	0.896
		10	0.453	0.673	38.188	9.977	0.431	0.693	0.871	0.896
	max	1	0.452	0.669	38.383	10.050	0.428	0.729	0.871	0.897
		2	0.453	0.672	38.705	10.028	0.428	0.730	0.872	0.897
		3	0.453	0.673	38.895	10.046	0.428	0.731	0.871	0.897
		4	0.453	0.672	38.958	10.087	0.427	0.730	0.871	0.898
		5	0.453	0.671	38.972	10.112	0.427	0.730	0.871	0.897
		6	0.453	0.672	39.066	10.097	0.426	0.730	0.871	0.898
		7	0.453	0.672	39.080	10.077	0.426	0.730	0.871	0.898
		8	0.452	0.672	39.108	10.085	0.426	0.730	0.871	0.898
		9	0.452	0.671	39.273	10.088	0.426	0.730	0.871	0.898
		10	0.451	0.671	39.288	10.098	0.426	0.730	0.870	0.897
	avg	1	0.452	0.669	38.383	10.050	0.428	0.729	0.871	0.897
		2	0.453	0.673	38.471	9.968	0.431	0.687	0.873	0.898
		3	0.454	0.673	38.359	9.980	0.431	0.687	0.873	0.899
		4	0.454	0.672	38.307	10.018	0.432	0.687	0.873	0.899
		5	0.454	0.672	38.309	10.016	0.432	0.686	0.873	0.899
		6	0.454	0.672	38.307	10.020	0.432	0.686	0.873	0.899
		7	0.454	0.672	38.295	9.988	0.432	0.687	0.873	0.900
		8	0.454	0.672	38.292	9.993	0.431	0.687	0.873	0.900
		9	0.454	0.672	38.285	10.011	0.432	0.687	0.873	0.899
		10	0.454	0.672	38.296	9.995	0.431	0.687	0.873	0.899

Table D7: Performance of semantic entailment depending on axioms, aggregation strategy and number of models for DeepGOGATMF-SE-Pred model