NEURAL INFORMATION PROCESSING SYSTEMS

Automatic Auxiliary Task Selection and Adaptive Weighting Boost Molecular Property Prediction

Zhiqiang Zhong[1,2,3], Davide Mottin[1]

[1]*Department of Computer Science, Aarhus University*
[2]*Institute for Advanced Studies, University of Luxembourg*
[3]*Faculty of Science, Technology and Medicine, University of Luxembourg*

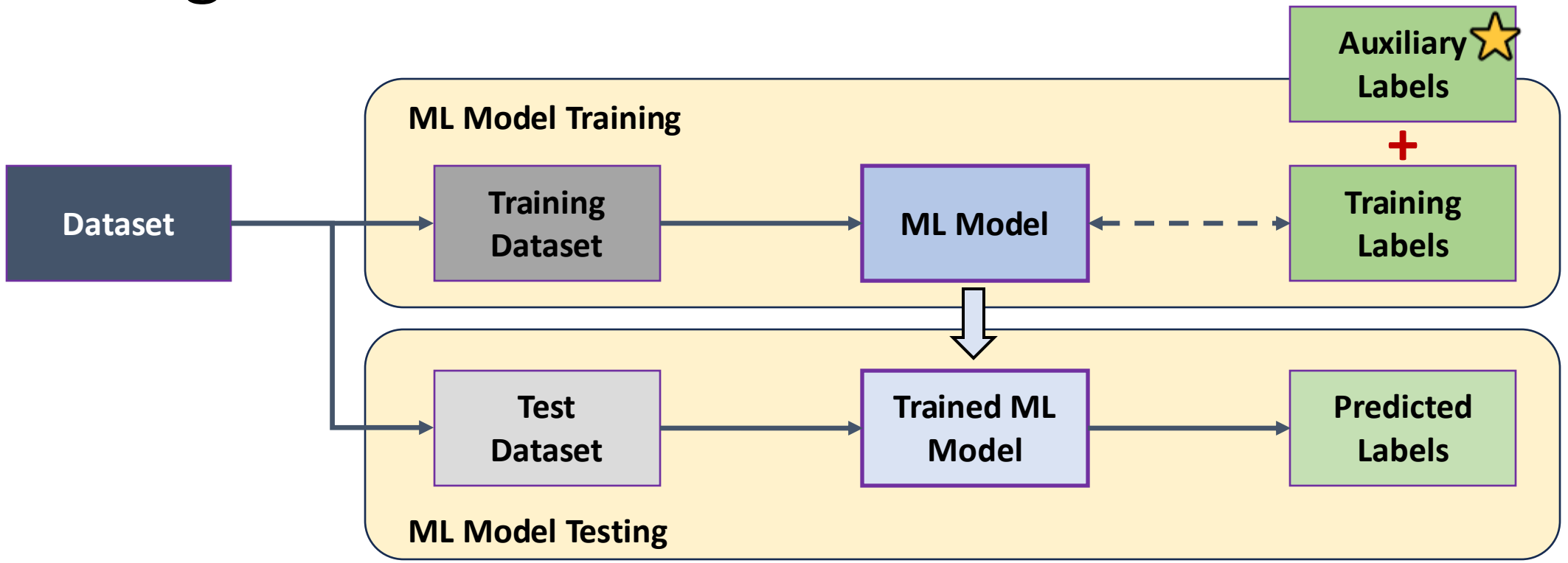
INSTITUTE FOR ADVANCED STUDIES (IAS)

Co-funded by the European Union

AARHUS UNIVERSITY

YOUNG INTERNATIONAL ACADEMICS

Background



- Introducing *auxiliary learning tasks* has been shown to be an effective way to address the problem of limited training labels in many applications.
- For example, incorporating chemical, physical, and toxicological profiles as auxiliary tasks helps ML models capture underlying data structures more effectively and improve generalization.

Motivation

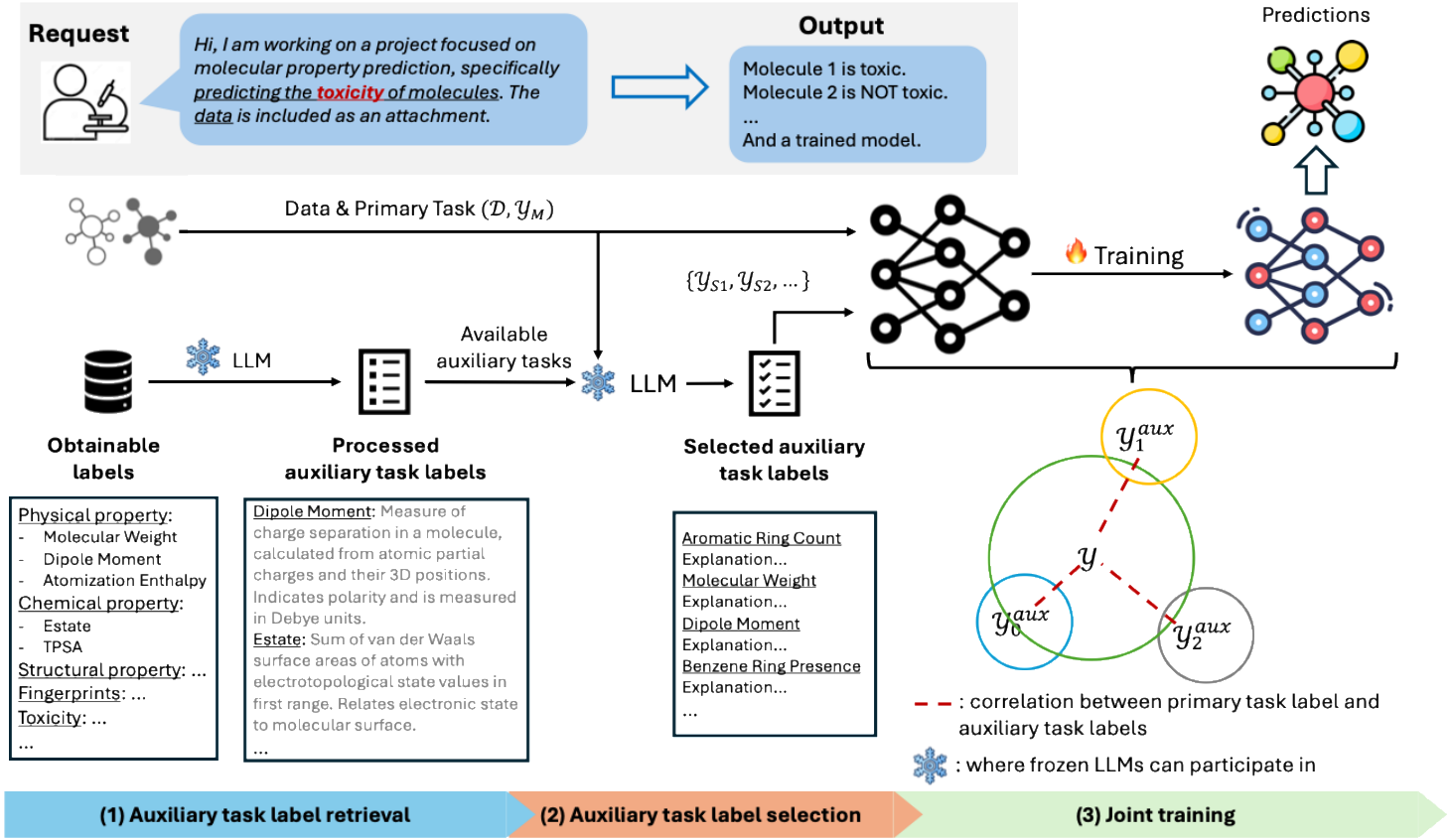
Model	Automatic Task Retrieval	Automatic Task Selection	Adaptive Weighting	Self-contained	Necessary Inputs
Unweighted Averages	✘	✘	✘	✔	Auxiliary Tasks
TAG [8]	✘	✘	✔	✔	Auxiliary Tasks
Task2Vec [1]	✘	✘	✔	✔	Auxiliary Tasks & Descriptions
MTDNN [25]	✘	✘	✘	✘	Relevant Auxiliary Datasets & Descriptions
Gradnorm [3]	✘	✘	✔	✔	Auxiliary Tasks
GS-Meta [68]	✘	✔	✘	✘	Auxiliary Tasks & Relation Graph
MolGroup [17]	✘	✘	✔	✘	Relevant Auxiliary Datasets & Descriptions
InstructMol [51]	✘	✘	✔	✘	Relevant Auxiliary Datasets & Descriptions
AutAuT (Ours)	✔	✔	✔	✔	None

- However, constructing high-quality auxiliary tasks is often complex and resource-intensive.
- In fields such as biology and chemistry, these challenges are further amplified by the scarcity, high cost, and time demands of obtaining domain-specific knowledge.

RQ: Can we automate the process of auxiliary task–enhanced ML model training?

Method - AutAuT

TL;DR: A fully automated framework that uses LLMs to retrieve *Auxiliary Tasks* and employs *Adaptive Weighting* to integrate them for molecular property prediction.



1) Auxiliary Task Retrieval

1) *Instruction*: Provides general guidance to the LLM, specifying its role in the retrieval process.
2) *Message*: A direct and clear request for the LLM to identify potential auxiliary task labels based on the given context

2) Auxiliary Task Selection

1) Search available information about these properties and write a brief summary...
2) Assess the relevance between retrieved properties and the primary task...
3) For the primary task, recommend K properties as auxiliary tasks...

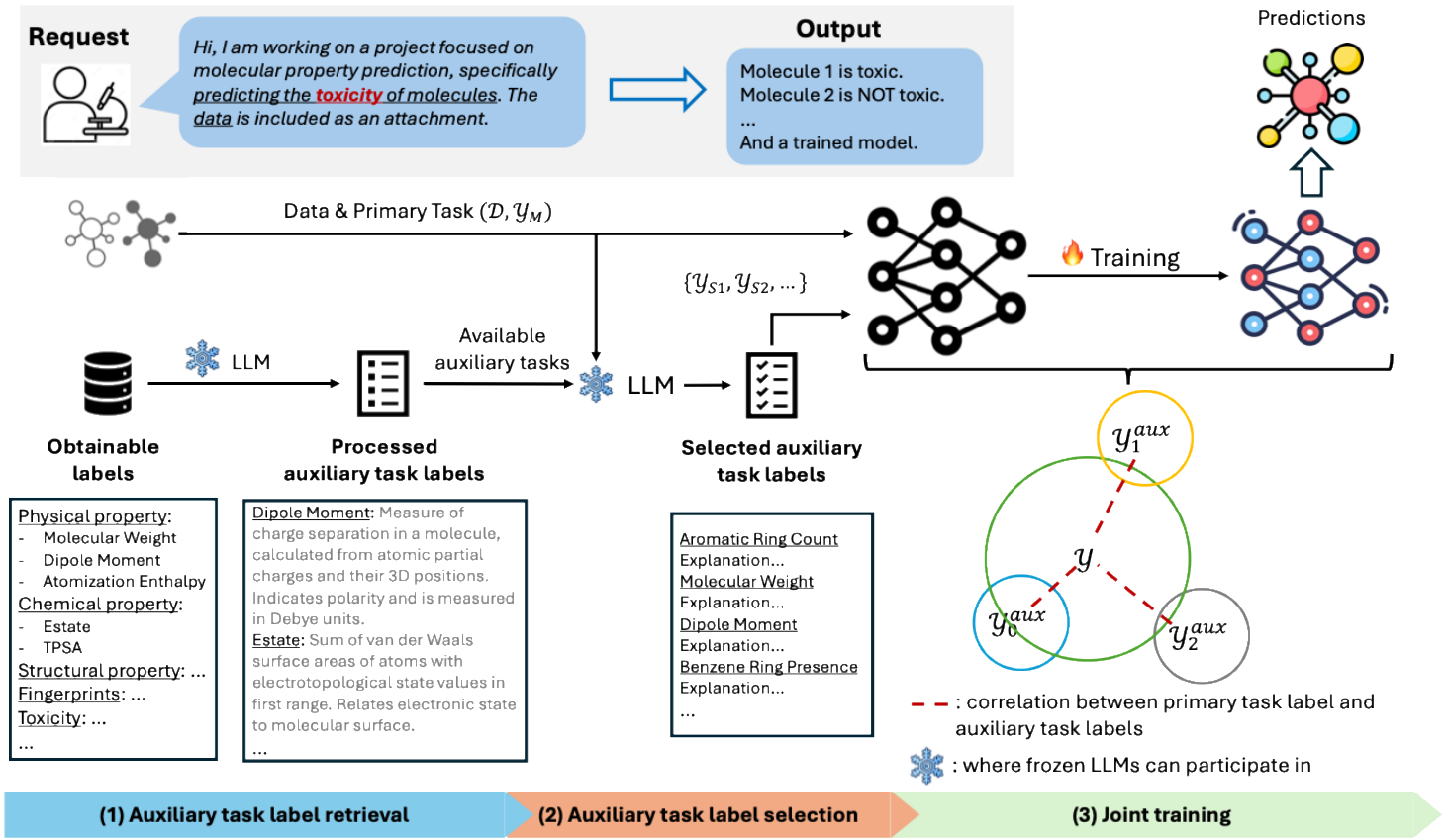
Method - AutAuT

Name	Category	Brief Description
Molecular Weight	Constitutional	Total mass of a molecule calculated as the sum of atomic weights of all atoms, providing fundamental information about molecular size and mass distribution.
Heavy Atom Molecular Weight	Constitutional	Sum of atomic weights of all non-hydrogen atoms in the molecule, useful for comparing core molecular frameworks.
Number of Valence Electrons	Constitutional	Total number of electrons in the outer shells of all atoms, crucial for understanding chemical bonding and reactivity patterns.
Total Formal Charge	Constitutional	Sum of all formal charges on atoms in the molecule, indicating overall molecular charge state and ionic character.
Topological Polar Surface Area	Topological	Sum of surfaces of all polar atoms (mainly oxygen and nitrogen), correlating with drug absorption, including intestinal absorption and blood-brain barrier penetration.
Labute Approximate Surface Area	Topological	Approximate molecular surface area calculated using Labute's method, useful for predicting physical properties and molecular interactions.
Balaban J Index	Topological	Topological index based on molecular connectivity, indicating molecular branching and cyclicity. Higher values suggest more branched structures.
Bertz Complexity	Topological	Measure of molecular complexity considering both size and branching patterns. Higher values indicate more complex molecular structures.
LogP	Electronic	Logarithm of octanol-water partition coefficient, predicting molecular lipophilicity and membrane permeability. Key for drug absorption.
Molar Refractivity	Electronic	Measure of total polarizability of a molecule, related to molecular volume and electronic properties. Important for predicting optical behavior.
EState VSA1	Electronic	Sum of van der Waals surface areas of atoms with electrotopological state values in first range. Relates electronic state to molecular surface.

1) Auxiliary Task Retrieval
   1) *Instruction*: Provides general guidance to the LLM, specifying its role in the retrieval process.
   2) *Message*: A direct and clear request for the LLM to identify potential auxiliary task labels based on the given context
2) Auxiliary Task Selection
   1) Search available information about these properties and write a brief summary…
   2) Assess the relevance between retrieved properties and the primary task…
   3) For the primary task, recommend K properties as auxiliary tasks…

Method - AutAuT

TL;DR: A fully automated framework that uses LLMs to retrieve *Auxiliary Tasks* and employs *Adaptive Weighting* to integrate them for molecular property prediction.



1) Auxiliary Task Retrieval
2) Auxiliary Task Selection
3) Learning to Learn from Selected Auxiliary Tasks
   1) *Weigh initialization*: the auxiliary task weights are initialized based on affinity scores.
   2) *Dynamic weight adaptation*: the auxiliary task weights are iteratively updated using gradient alignment and validation performance.
   3) *Dynamic weight adaptation*: the ML model parameters are optimized with fixed auxiliary task weights.

Theorem 1

$$\|\nabla \mathcal{L}_{\mathrm{M}}(\theta) - \sum_{k=1}^{K} \alpha_k \nabla \mathcal{L}_{\mathrm{S}}^{k}(\theta)\| = 0.$$

(Subspace Alignment for Optimization)

Theorem 2

$$\mathcal{E}(\theta) \leq \hat{\mathcal{E}}(\theta) + c \cdot \mathcal{R}_n(\mathcal{H}_{\boldsymbol{\alpha}}),$$

(Generalization Bounds with Auxiliary Tasks)

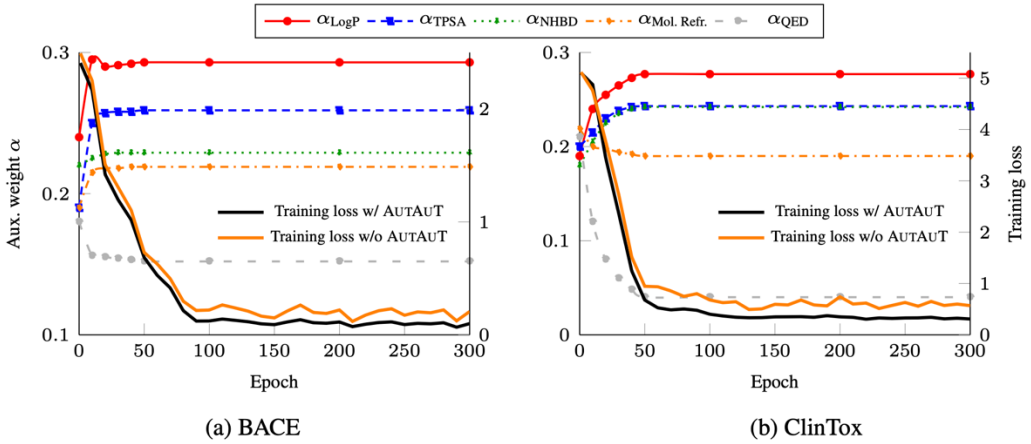
Experiments

	Classification (ROC-AUC % ↑)						Regression (RMSE ↓)		
Datasets	**BBBP**	**BACE**	**ClinTox**	**Tox21**	**ToxCast**	**SIDER**	**ESOL**	**FreeSolv**	**Lipo**
# Molucules	2,039	1,513	1,478	7,831	8,575	1,427	1,128	642	4,200
# Tasks	1	1	2	12	617	27	1	1	1
GCN	63.45±0.05	74.83±0.18	56.28±0.09	74.63±0.06	65.38±0.30	62.24±0.27	3.165±0.007	3.752±0.013	1.672±0.008
+Beam search	66.08±0.02	68.42±0.03	58.32±0.08	75.30±0.01	66.45±0.22	62.87±0.25	3.205±0.020	3.415±0.018	1.685±0.012
+TAG	64.65±0.02	72.10±0.02	57.12±0.06	71.78±0.01	65.08±0.21	62.63±0.23	3.340±0.035	3.885±0.055	1.722±0.014
+Task2vec	68.25±0.01	75.38±0.03	49.55±0.07	70.20±0.01	63.62±0.20	62.25±0.23	3.278±0.042	3.437±0.030	1.742±0.016
+MTDNN	66.63±0.02	70.12±0.03	53.25±0.08	71.95±0.01	64.53±0.25	61.95±0.26	3.172±0.055	3.420±0.040	1.715±0.015
+UA	60.38±0.01	62.48±0.03	51.95±0.08	70.95±0.01	62.75±0.24	60.18±0.22	3.428±0.072	4.210±0.060	1.810±0.017
+Gradnorm	61.55±0.01	65.12±0.04	53.45±0.07	59.88±0.05	62.05±0.23	59.20±0.22	3.558±0.082	4.410±0.065	1.855±0.018
+PF	57.12±0.03	65.22±0.05	56.45±0.04	51.12±0.02	60.88±0.21	58.62±0.20	3.658±0.075	4.508±0.072	1.902±0.021
+MolGroup	68.42±0.02	77.68±0.02	60.18±0.03	76.12±0.01	66.95±0.23	63.12±0.25	3.125±0.025	3.245±0.020	1.702±0.012
+GS-Meta	66.72±0.02	75.71±0.41	63.23±0.06	75.23±0.08	65.94±0.21	63.22±0.60	3.035±0.013	3.444±0.061	1.839±0.027
+InstructMol	68.50±0.04	80.21±0.18	64.02±0.05	76.69±0.04	66.94±0.48	64.30±0.72	2.612±0.034	2.108±0.046	1.252±0.018
+AutAuT	**69.72**±0.01	**83.10**±0.02	**81.00**±0.03	**77.55**±0.01	**68.05**±0.19	**64.62**±0.24	**1.885**±0.028	**1.975**±0.022	**1.285**±0.010
GIN	66.82±0.09	77.45±0.21	56.48±0.18	75.32±0.17	62.35±0.05	60.25±0.20	2.895±0.012	3.865±0.018	1.692±0.011
+Beam search	67.78±0.02	81.92±0.04	76.95±0.12	77.05±0.02	65.28±0.22	62.57±0.22	2.943±0.028	3.375±0.022	1.688±0.014
+TAG	61.25±0.02	72.15±0.03	58.25±0.06	72.08±0.01	64.38±0.21	61.75±0.21	3.275±0.060	4.002±0.048	1.752±0.015
+Task2vec	68.48±0.01	75.10±0.03	48.62±0.04	69.12±0.01	63.62±0.24	60.38±0.26	3.232±0.045	3.452±0.032	1.775±0.017
+MTDNN	66.92±0.03	70.35±0.03	53.28±0.07	72.32±0.01	64.58±0.25	61.92±0.26	3.182±0.052	3.468±0.042	1.722±0.014
+UA	60.75±0.01	62.18±0.03	52.18±0.08	71.22±0.01	62.58±0.23	59.82±0.22	3.442±0.065	4.315±0.053	1.822±0.019
+Gradnorm	61.68±0.02	65.38±0.04	53.58±0.08	60.25±0.05	62.08±0.23	58.55±0.23	3.602±0.080	4.508±0.062	1.878±0.020
+PF	57.38±0.03	65.58±0.05	56.38±0.04	51.38±0.02	60.92±0.22	58.38±0.21	3.695±0.078	4.608±0.070	1.932±0.021
+MolGroup	69.88±0.01	83.12±0.02	81.25±0.04	77.72±0.01	67.55±0.22	64.28±0.24	1.882±0.030	1.952±0.025	1.288±0.011
+GS-Meta	66.46±0.09	79.46±0.00	73.54±0.12	76.27±0.08	66.14±0.31	62.96±0.25	2.060±0.012	3.848±0.024	1.434±0.041
+InstructMol	69.87±0.02	81.99±0.04	76.16±0.06	77.91±0.02	68.44±0.14	64.25±0.62	1.878±0.019	1.960±0.022	1.289±0.026
+AutAuT	**70.58**±0.01	**83.62**±0.02	**82.52**±0.04	**78.32**±0.01	**69.38**±0.22	**65.08**±0.25	**1.870**±0.025	**1.900**±0.020	**1.270**±0.010
Graphormer	67.05±0.04	79.18±0.13	78.42±0.12	75.55±0.18	67.12±0.05	70.08±0.21	2.102±0.012	1.795±0.010	1.282±0.010
+Beam search	67.78±0.03	81.12±0.03	78.22±0.09	77.08±0.01	67.82±0.21	70.55±0.24	2.225±0.040	1.825±0.015	1.322±0.014
+TAG	61.52±0.02	72.25±0.04	59.35±0.07	72.55±0.01	66.25±0.23	69.25±0.25	2.285±0.060	1.865±0.020	1.375±0.018
+Task2vec	68.52±0.01	75.18±0.04	48.85±0.03	69.25±0.01	65.12±0.22	67.55±0.27	2.275±0.040	1.852±0.020	1.362±0.015
+MTDNN	67.08±0.02	70.52±0.03	53.52±0.07	72.45±0.01	65.55±0.25	68.12±0.26	2.202±0.042	1.875±0.018	1.335±0.012
+UA	60.82±0.01	62.22±0.03	52.52±0.08	71.52±0.01	63.52±0.22	66.72±0.24	2.455±0.050	2.055±0.020	1.432±0.016
+Gradnorm	62.02±0.02	65.52±0.04	53.82±0.07	60.52±0.05	62.82±0.23	65.52±0.25	2.605±0.070	2.205±0.040	1.525±0.020
+PF	57.52±0.03	65.82±0.05	56.52±0.04	51.52±0.02	61.22±0.22	64.82±0.23	2.655±0.080	2.305±0.050	1.625±0.018
+MolGroup	69.72±0.01	83.15±0.02	81.52±0.04	77.75±0.01	68.85±0.22	70.55±0.27	1.902±0.030	1.805±0.010	1.292±0.010
+GS-Meta	68.12±0.00	81.06±0.10	80.00±0.11	77.29±0.09	67.90±0.20	70.14±0.25	2.064±0.035	1.881±0.016	1.330±0.019
+InstructMol	68.99 ±0.02	83.52±0.02	81.41±0.02	78.05±0.04	69.03±0.07	70.49±0.23	1.901±0.014	1.811±0.002	1.230±0.009
+AutAuT	**70.32**±0.01	**83.72**±0.02	**82.55**±0.04	**78.52**±0.01	**69.38**±0.21	**71.55**±0.26	**1.872**±0.025	**1.772**±0.015	**1.262**±0.010

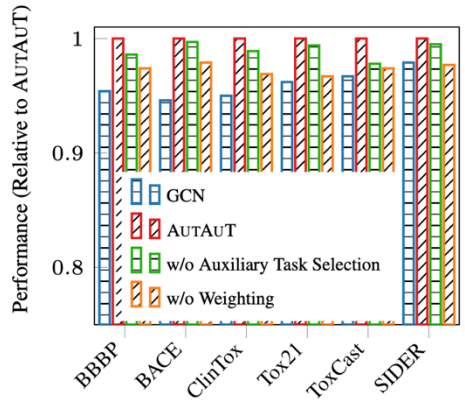
	Classification (ROC-AUC % ↑)						Regression (RMSE ↓)		
Datasets	**BBBP**	**BACE**	**ClinTox**	**Tox21**	**ToxCast**	**SIDER**	**ESOL**	**FreeSolv**	**Lipo**
GCN	63.45±0.05	74.83±0.18	56.28±0.09	74.63±0.06	65.38±0.30	62.24±0.27	3.165±0.007	3.752±0.013	1.672±0.008
GIN	66.82±0.09	77.45±0.21	56.48±0.18	75.32±0.17	62.35±0.05	60.25±0.20	2.895±0.012	3.865±0.018	1.692±0.011
Graphormer	67.05±0.04	79.18±0.13	78.42±0.12	75.55±0.18	67.12±0.05	70.08±0.21	2.102±0.012	1.795±0.010	1.282±0.010
D-MPNN	70.83±0.42	81.19±0.51	90.95±0.63	76.24±0.47	64.97±0.24	57.31±0.45	1.045±0.070	2.080±0.090	0.685±0.015
AttentionFP	64.15±1.48	78.24±0.31	80.51±0.26	77.00±0.35	63.52±0.34	60.12±6.03	0.880±0.025	2.075±0.150	0.720±0.002
MGCN	65.27±0.43	73.01±0.61	89.51±1.22	77.15±0.50	66.03±0.48	58.02±1.93	1.100±0.060	2.800±0.100	0.730±0.012
N-Gram	70.15±0.52	78.09±1.37	75.95±2.97	77.21±0.45	62.04±0.60	58.94±5.50	1.150±0.045	2.850±0.110	0.735±0.013
PretrainGNN	68.91±0.98	83.97±0.64	72.95±0.61	77.45±0.38	63.01±0.87	62.03±5.94	1.100±0.050	2.750±0.010	0.740±0.004
GPT-GNN	63.96±1.23	71.98±0.67	65.04±1.00	77.12±0.60	62.00±0.45	58.47±4.00	1.200±0.070	3.000±0.150	0.800±0.020
GROVER$_{base}$	65.02±0.19	81.01±0.29	74.92±0.60	77.31±0.41	62.97±0.70	61.01±5.50	1.085±0.075	2.270±0.050	0.820±0.012
GROVER$_{large}$	60.04±0.31	80.99±0.36	74.01±0.55	77.10±0.45	63.99±0.50	60.02±6.00	0.980±0.080	2.180±0.045	0.815±0.009
3D-Infomax	67.02±1.23	84.01±1.31	78.03±1.43	77.25±0.32	65.96±0.80	62.04±5.90	0.950±0.030	2.600±0.080	0.800±0.016
GraphMVP	71.98±1.48	83.02±0.79	85.03±0.90	77.40±0.25	65.94±0.60	63.01±5.30	0.890±0.030	2.330±0.250	0.830±0.006
MolCLR	71.51±1.96	83.48±0.52	88.97±1.10	77.35±0.34	66.98±0.70	64.05±5.70	0.940±0.120	2.580±0.220	0.650±0.007
Uni-Mol	72.83±0.48	**86.52**±0.39	89.97±1.00	77.50±0.22	67.03±0.60	65.94±5.50	0.785±0.100	1.620±0.050	0.605±0.011
GEM	72.21±0.34	85.49±0.97	91.98±1.45	77.55±0.15	66.98±0.41	66.97±0.48	0.800±0.025	1.870±0.090	0.660±0.010
Pin-Tuning	70.87±0.61	81.21±0.20	89.81±1.44	77.20±0.27	66.92±0.98	66.36±0.81	0.845±0.204	1.935±0.104	0.673±0.112
LAC	72.44±0.95	83.47±0.11	92.11±1.19	77.66±0.61	67.01±0.34	67.02±0.65	0.794±0.100	1.827±0.121	0.657±0.108
AutAuT	**73.04**±0.58	86.03±1.14	**92.31**±1.23	**78.20**±0.12	**67.54**±0.76	**67.53**±0.82	**0.760**±0.035	**1.600**±0.040	**0.580**±0.012

- AutAuT surpasses 10 auxiliary task selection methods across 9 molecular property prediction tasks.
- AutAuT demonstrates superior molecular property prediction performance against 10 state-of-the-art molecular property prediction methods.

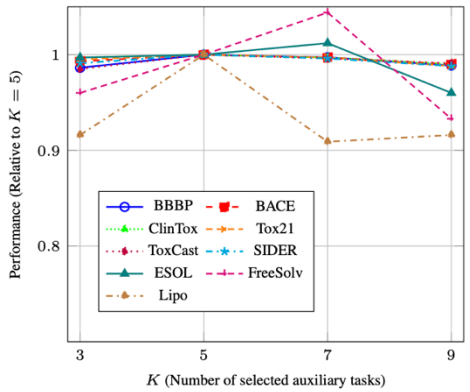
Experiments



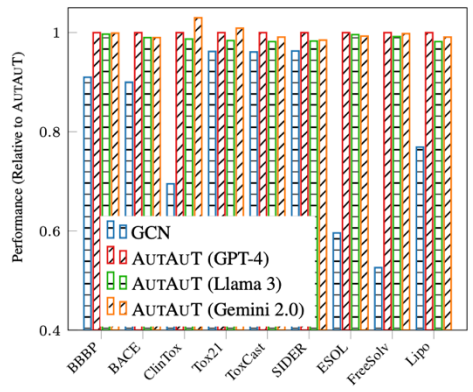
A



B



C



D

A: Auxiliart task influence is correctly inferred.

B: Auxiliary task selection and adaptive weighting complement each other.

C: The number of selected auxiliary tasks (K) balance informativeness and noise.

D: All major LLMs produce useful task suggestions and improve over baseline.

Thank you!

Contact

Paper