Transfer Learning for Prediction of Liquid-Liquid Phase Separation

Karlo Barjaktarić[1], Patrik Paunović[1], Leonard Lečić[1], Erik Otović[1], Joana Calvário[2], Ana Sofia Pina[2], Yoshimasa Kawaguchi[3], Daniela Kalafatovic[1], Goran Mauša*[1]

DOI: https://doi.org/10.17952/38EPS.2026.P250

*Contact email: goran.mausa@uniri.hr

[1] University of Rijeka, Faculty of Engineering, Rijeka, Croatia
[2] Institute of Chemical and Biological Technology, NOVA University of Lisbon, Oeiras, Portugal
[3] Institute for Chemical Research, Kyoto University, Kyoto, Japan

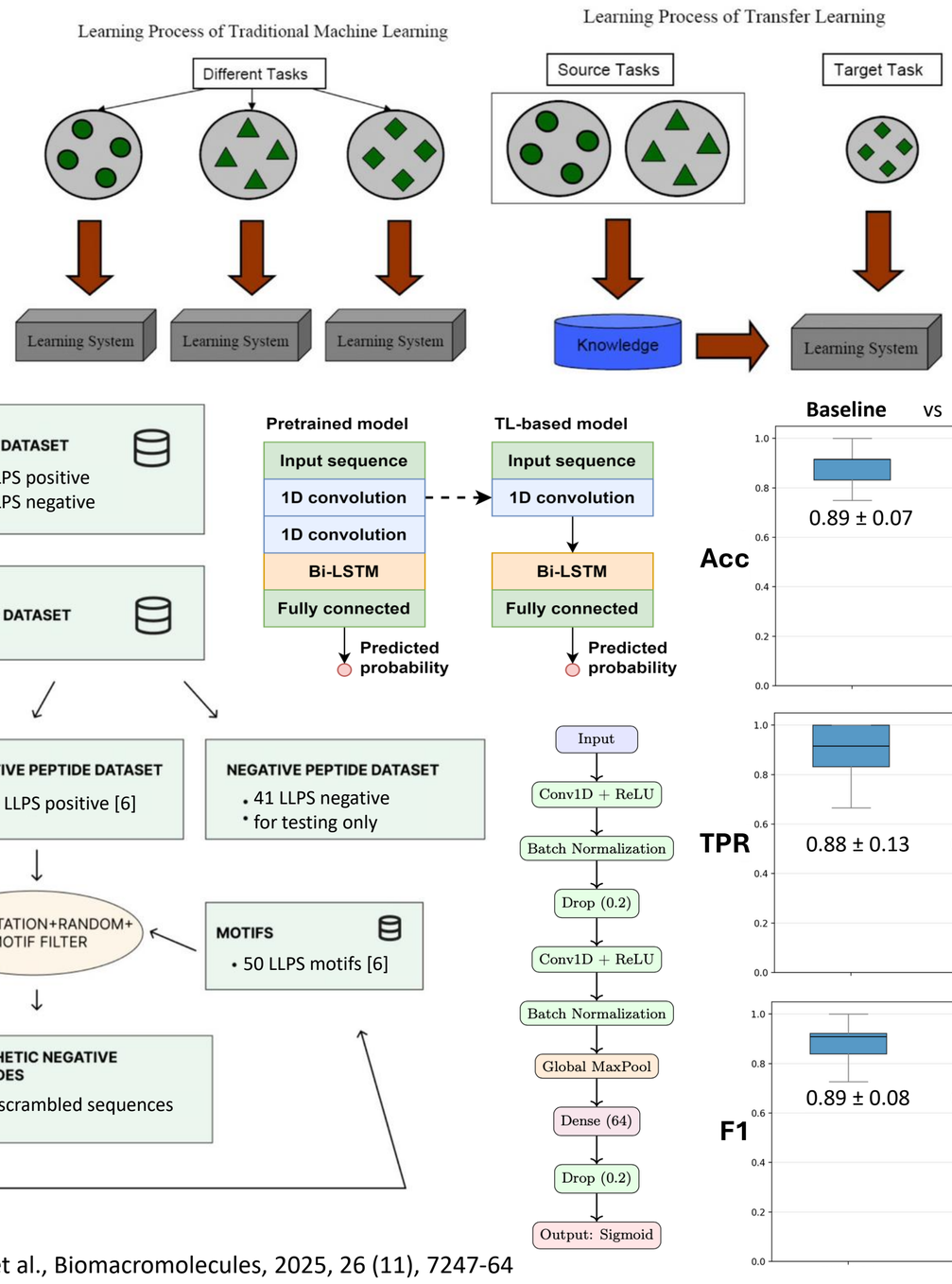
1. Amount of Data Required to Train a Peptide Classifier



	Catalytic [1]	Self-assembly [2]	AVP [3]	Permeability [4]	AMP [3, 5]
Size:	~100	~400	~1,000	~7,000	~10,000
Ratio:	68 : 31	67 : 33	61 : 39	69: 31	56 : 44
Acc:	95%	82%	85%	82 %	92%
AUC:	0.95	0.86	0.88	0.78	0.97

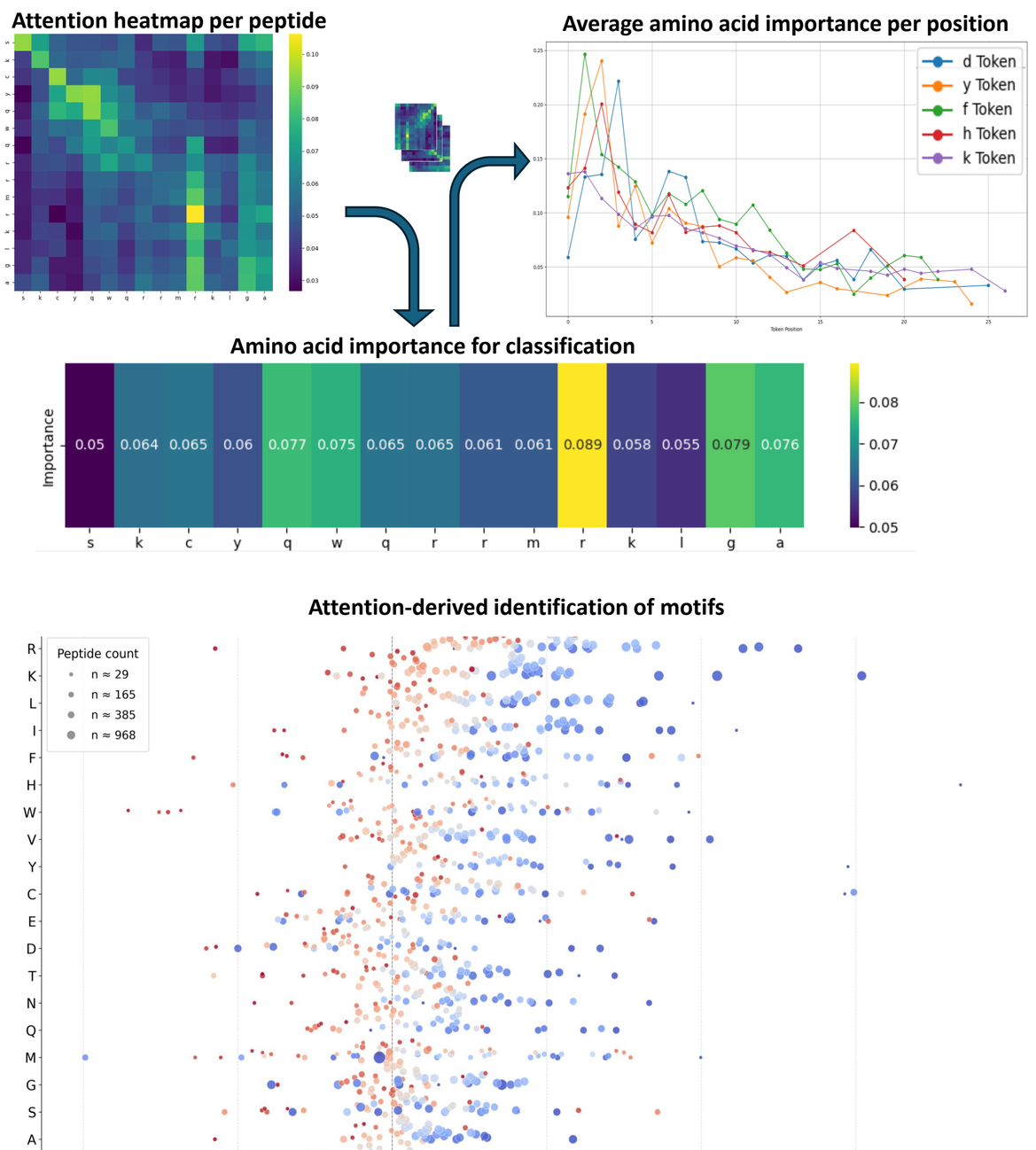
[1] Mauša et al., MRS Advances, 2023, 8, 1068–1074
[2] Njirjak et al., Nature Machine Intelligence, 2024, 6, 1487–1500
[3] Otović et al., Journal of Chemical Information and Modeling, 2022, 62 (12), 2961-2972
[4] Cabezón Vizoso et al., Digital Discovery, 2025, 4, 1259-1275
[5] Negovetić et al., Digital Discovery, 2024, 3, 1182

2. Protein – to – Peptide Transfer Learning for LLPS Prediction



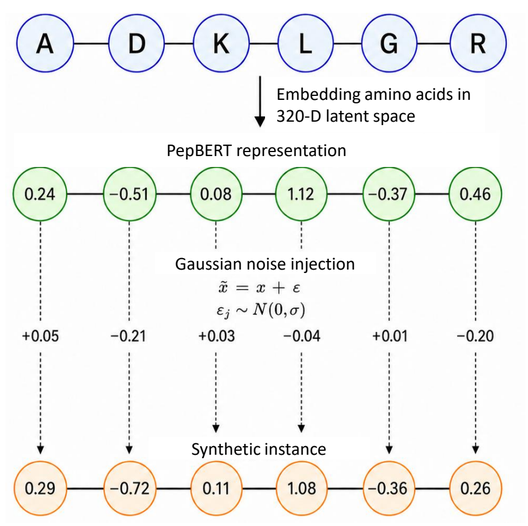
[6] Calvário et al., Biomacromolecules, 2025, 26 (11), 7247-64

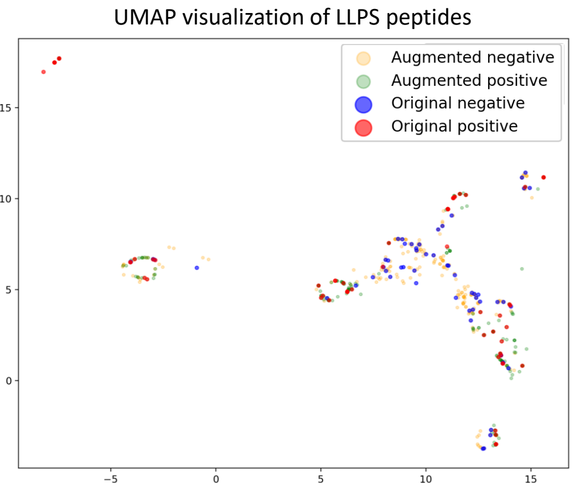
3. Position-specific Sequence Determinants for Peptide Classification Derived from Token Importance



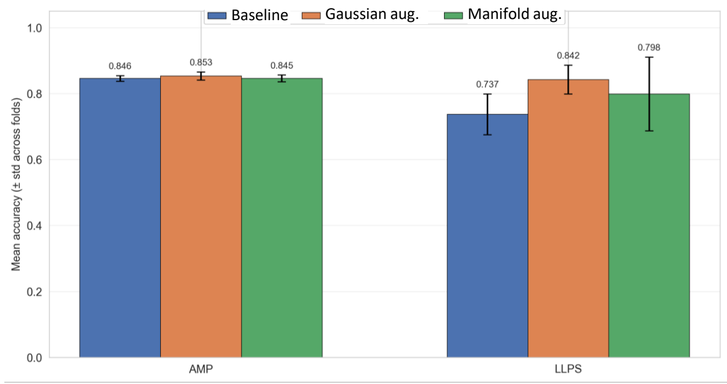
[7] Du et al., Future Foods, 2026, 13, 2961-2972
[8] Devlin et al., Proceedings of NAACL-HLT 2019, 4171--4186

4. Noise Injection-based Augmentation Strategy





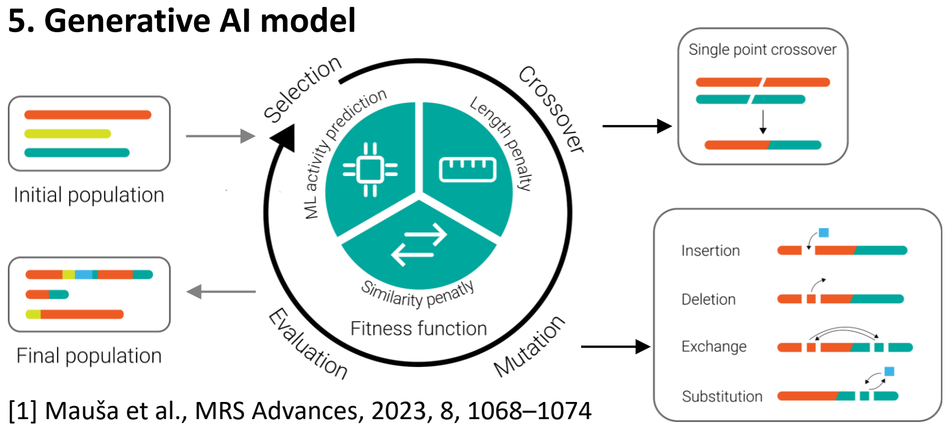
Data	Representation	Acc	F1	MCC
AMP	PepBERT (320-D) [7]	**84.6**	**0.80**	**0.68**
AMP	SeqProps (28-D) [3]	79.6	0.73	0.57
LLPS	PepBERT (320-D) [7]	73.7	0.74	0.48
LLPS	SeqProps (28-D) [3]	**84.2**	**0.84**	**0.69**

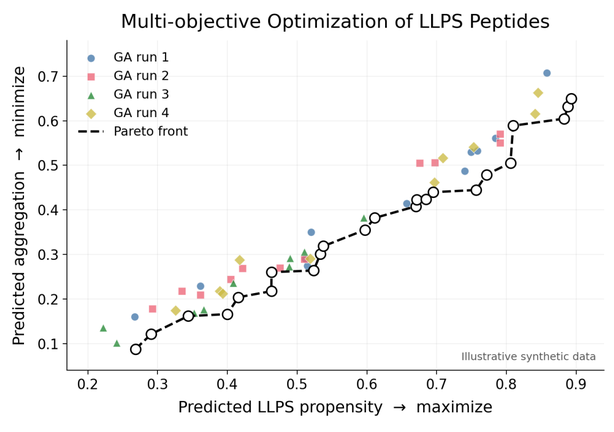


Settings	Acc	F1	MCC
PepBERT + Gaussian injection (noise =0.5, fraction = 100)	84.2	0.85	0.70
SeqProps + Gaussian injection (noise =0.1, fraction = 100)	86.0	0.86	0.72

[3] Otović et al., Journal of Chemical Information and Modeling, 2022, 62 (12), 2961-2972
[7] Du et al., Future Foods, 2026, 13, 2961-2972

5. Generative AI model





[1] Mauša et al., MRS Advances, 2023, 8, 1068–1074

Acknowledgment

NextGenerationEU project: SCI-PHI (uniri-mzi-25-25)

Funded by the European Union NextGenerationEU

Croatian Science Foundation projects: Super Deshpet 2.0 (IP-2025-02-3357) MOPED (IP-2026-4962)

HrZZ

COST action: SNOOPY (CA23111)

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

SNOOPY