Result No.	Paper Name	Link	Related	Code Available	Comments	Purpose		For Related, Please rate one these 3		
Tushar								Yes	No	Maybe
1	Rethinking Positional Encoding	https://arxiv.org/abs/2107.02561	No				**Count for each**	14	42	12
2	A Deep Learning-Based Particle-in-Cell Method for Plasma Simulations	https://arxiv.org/abs/2107.02232	No							
3	TransformerFusion: Monocular RGB Scene Reconstruction using Transformers	https://arxiv.org/abs/2107.02191	No							
4	Imputation-Free Learning from Incomplete Observations	https://arxiv.org/abs/2107.01983	No							
5	A contextual analysis of multi-layer perceptron models in classifying hand-written digits and letters: limited resources	https://arxiv.org/abs/2107.01782	Maybe		Techniques used in this paper, can be helpful to increase the performance of MLP					
6	Short-term probabilistic photovoltaic power forecast based on deep convolutional long short-term memory network and kern	https://arxiv.org/abs/2107.01343	No							
7	Global Filter Networks for Image Classification	https://arxiv.org/abs/2107.00645	Maybe	Yes	(*) Perhaps interesting if order awareness becomes relevant for textual tokens; -- Improvement on Vision Transformers					
8	Machine Learning and Deep Learning for Fixed-Text Keystroke Dynamics	https://arxiv.org/abs/2107.00507	No							
9	Dep-L0: Improving L0-based Network Sparsification via Dependency Modeling	https://arxiv.org/abs/2107.00070	No							
10	Rethinking Token-Mixing MLP for MLP-based Vision Backbone	https://arxiv.org/abs/2106.14882	Yes		Improvement on Token Mixing MLP					
Andor										
11	A 3D CNN Network with BERT For Automatic COVID-19 Diagnosis From CT-Scan Images	https://arxiv.org/abs/2106.14403	No		Task specific, no novel MLP use					
12	Generalized Zero-Shot Learning using Multimodal Variational Auto-Encoder with Semantic Concepts	https://arxiv.org/abs/2106.14082	No		Focus on VAE-s					
13	Exploring Corruption Robustness: Inductive Biases in Vision Transformers and MLP-Mixers	https://arxiv.org/abs/2106.13122	Maybe	can reference in report	Transformers are more robust to corruption than CNN and MLP mixers					
14	Handwritten Digit Recognition using Machine and Deep Learning Algorithms	https://arxiv.org/abs/2106.12614	No		How is this even a paper in 2021? (SVM MLP CNN on mnist...)					
15	Vision Permutator: A Permutable MLP-Like Architecture for Visual Recognition	https://arxiv.org/abs/2106.12368	Yes	yes	Vision specific MLP mixer, retains height&width dimension of patches, might give ideas					
16	Regularization is all you Need: Simple Neural Nets can Excel on Tabular Data	https://arxiv.org/abs/2106.11189	Yes		Might give useful modern regularization ideas					
17	α-Stable convergence of heavy-tailed infinitely-wide neural networks	https://arxiv.org/abs/2106.11064	No		Math heavy with proofs, seems strictly theoretical					
18	Exoskeleton-Based Multimodal Action and Movement Recognition: Identifying and Developing the Optimal Boosted Learning Approa	https://arxiv.org/abs/2106.10331	No		Task specific comparions, k-NN works best					
19	Predicting gender of Brazilian names using deep learning	https://arxiv.org/abs/2106.10156	No		Basic binary classification, MLP, RNN, CNN comparison					
20	Learning and Meshing from Deep Implicit Surface Networks Using an Efficient Implementation of Analytic Marching	https://arxiv.org/abs/2106.10031	No		Very specific graphics task					
Tim										
21	Early fault detection with multi-target neural networks	https://arxiv.org/abs/2106.08957	No		no new MLP approach, multi-target MLPs detect faults earlier than single-target MLPs					
22	Distilling Self-Knowledge From Contrastive Links...	https://arxiv.org/abs/2106.08541	No		use basic MLP, no new MLP approach					
23	PairConnect: A Compute-Efficient MLP Alternative to Attention	https://arxiv.org/abs/2106.08235	Yes	no	(*) for order aware models, perhaps for the literature; models the pairwise interaction between words by explicit pairwise word embeddings					
24	Evaluating Modules in Graph Contrastive Learning	https://arxiv.org/abs/2106.08171	No		framework for graph contrastive learning					
25	MLP Singer: Towards Rapid Parallel Korean Singing Voice Synthesis	https://arxiv.org/abs/2106.07886	No		MLP-Mixer for attention-free image classification					
26	Grasping force estimation using state-space model and Kalman filter	https://arxiv.org/abs/2106.07639	No		math heavy, comparing new approach based on the state-space with Kalman filter with MLP					
27	S2-MLP: Spatial-Shift MLP Architecture for Vision	https://arxiv.org/abs/2106.07477	Yes	code on PaddlePaddle	new MLP architecture, channel-mixing MLP instead of MLP-Mixer					
28	Pressure-Informed Velocity Estimation in a Subsonic Jet	https://arxiv.org/abs/2106.07110	No		use basic MLP for specific task, nothing new					
29	AI Empowered Resource Management for Future Wireless Networks	https://arxiv.org/abs/2106.06178	No		theoretical, show that GNNs to MLPs					
30	Empirical observations on the effects of data transformation...	https://arxiv.org/abs/2106.05855	No		comparing different classifiers for geozone classification					
Bhakti										
31	Graph-MLP: Node Classification without Message Passing in Graph	https://arxiv.org/abs/2106.04051	Yes	Yes (Ansgar)	a novel MLP-based method, Graph-MLP for learning graph node feature distribution					
32	Adversarial Classification of the Attacks on Smart Grids Using Game Theory and Deep Learning	https://arxiv.org/abs/2106.03209	No	can refernce in report	use the MLPs instead of conventional state estimators					
33	Solving hybrid machine learning tasks by traversing weight space geodesics	https://arxiv.org/abs/2106.02793	No		a geometric framework					
34	Spline Positional Encoding for Learning 3D Implicit Signed Distance Fields	https://arxiv.org/abs/2106.01553	No		related to 3D shapes					
35	When Vision Transformers Outperform ResNets without Pretraining or Strong Data Augmentations	https://arxiv.org/abs/2106.01548	No	referencing maybe	a detailed analysis of the convolution-free ViTs and MLP-Mixers from the lens of the loss landscape geometry					
36	Container: Context Aggregation Network	https://arxiv.org/abs/2106.01401	No		shown that disparate architectures such as Transformers, depth-wise CNNs and MLP-based methods are closely related via an affinity matrix used for context aggregation					
37	Towards Deeper Deep Reinforcement Learning	https://arxiv.org/abs/2106.01151	No		Deep Reinforcement Learning					
38	One4all User Representation for Recommender Systems in E-commerce	https://arxiv.org/abs/2106.00573	Yes		shoppingBERT (compare the generalizability of two learning strategies, i.e., transfer learning through the proposed model, ShopperBERT, vs. learning from scratch)					
39	SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers	https://arxiv.org/abs/2105.15203	No	No (could be an idea for ot	(*) cool idea of segmenting text for tasks of sentiment analysis ; Light weight MLP decoder					
40	Can Attention Enable MLPs To Catch Up With CNNs?	https://arxiv.org/abs/2105.15078	Yes	referencing(more of survey	Comparision of different new architectures including MLP mixer					
Shirney										
41	same as 40									
42	A remark on a paper of Krotov and Hopfield	https://arxiv.org/abs/2105.15034	No		just a short remark					
43	Transformer-Based Deep Image Matching for Generalizable Person Re-identification	https://arxiv.org/abs/2105.14432	No		transformer-based					
44	Gradient-Free Neural Network Training via Synaptic-Level Reinforcement Learning	https://arxiv.org/abs/2105.14383	Maybe	No	an algorithm based on reinforcement learning (RL) to generate and apply a simple biologically-inspired synaptic-level learning policy for multi-layerperceptron (MLP) models.					
45	Less is More: Pay Less Attention in Vision Transformers	https://arxiv.org/abs/2105.14217	Yes	Yes	hierarchical Transformer where they use pure multi-layer perceptrons (MLPs)	image				
46	MixerGAN: An MLP-Based Architecture for Unpaired Image-to-Image Translation	https://arxiv.org/abs/2105.14110	Maybe	No	mixerGAN	image				
47	An In-Memory Analog Computing Co-Processor for Energy-Efficient CNN Inference on Mobile Devices	https://arxiv.org/abs/2105.13904	No							
48	Reduced Complexity Neural Network Equalizers for Two-dimensional Magnetic Recording	https://arxiv.org/abs/2105.13508	Maybe	No	four variants of reduced complexity MLP (RC-MLP) architectures	information areal density of hard disk drives				
49	Flow based features and validation metric for machine learning reconstruction of PIV data	https://arxiv.org/abs/2105.13429	No		Particle image velocimetry					
50	Structure and lattice thermal conductivity of grain boundaries in silicon by using machine learning potential and molecular dynamics	https://arxiv.org/abs/2105.11959	No		Grain boundaries (GBs) in polycrystalline sil-ico					
51	DTNN: Energy-efficient Inference with Dendrite Tree Inspired Neural Networks for Edge Vision Applications	https://arxiv.org/abs/2105.11848	No		Dendrite-Tree based Neural Network (DTNN)					
Franck										
51	PGDOT -- Perturbed Gradient Descent Adapted with Occupation Time	https://arxiv.org/abs/2005.04507	No	no	perturbed gradient descent (PGD), Adam, AMSGrad, RMSProp , MLP					
52	Mixture of ELM based experts with trainable gating network	https://arxiv.org/abs/2105.11706	No		Ensemble learning, Mixture of experts, Extreme learning machine, Neural network based ensemble learning					
53	Rotation Invariant Point Cloud Classification: Where Local Geometry Meets Global Topology	https://arxiv.org/abs/1911.00195	No							
54	MLP-Mixer: An all-MLP Architecture for Vision	https://arxiv.org/abs/2105.01601	Yes							
55	Deep Learning Methods for Vessel Trajectory Prediction based on Recurrent Neural Networks	https://arxiv.org/abs/2101.02486	Maybe	yes	RNN, LSTM, Deep Learning, Automatic Identificazion System (AIS), trajectory Prediction					
56	BGC: Multi-Agent Group Belief with Graph Clustering	https://arxiv.org/abs/2008.08808	No	No	Graph Clustering, KNN, Split Loss, Graph Attention Network					
57	Neural Network Potentials: A Concise Overview of Methods	https://arxiv.org/abs/2107.03727	No	no						
58	Enhancing Transformers with Gradient Boosted Decision Trees for NLI Fine-Tuning	https://arxiv.org/abs/2105.03791	No	no	Transfer Learning, Natural Language Inference (NLI), Gradient Boosted Decision Trees (GBDTs), MLP					
59	Machine Learning and Deep Learning for Fixed-Text Keystroke Dynamics	https://arxiv.org/abs/2107.00507	Maybe		MLP, XGBoost, Keystroke Dynamics, deep Learning techniques					
60	A Neural Network Perturbation	https://arxiv.org/abs/2009.03192	Yes		DNN, Neuronal Network, Partial derivatives of MLP					
Tim										
61	Beyond Self-attention: External Attention using Two Linear Layers for Visual Tasks	https://arxiv.org/abs/2105.02358	No		MLP for vision, that uses two external, small, learnable, shared memories, implemented simply using two cascaded linear layers and two normalization layers					
62	RepMLP: Re-parameterizing Convolutions into Fully-connected Layers for Image Recognition	https://arxiv.org/abs/2105.01883	Yes	yes	already referenced in overleaf, MLP-style neural network building block for image recognition					
	Pay Attention to MLPs	https://arxiv.org/abs/2105.08050	Yes							

Multi-label papers

	Paper Title	Link	Related	Code Available	Comments
1	Priberam at MESINESP Multi-label Classification of Medical Texts Task	https://arxiv.org/abs/2105.05614	Maybe	Yes	
2	Enhancing Label Correlation Feedback in Multi-Label Text Classification via Multi-Task Learning	https://arxiv.org/abs/2106.03103			1) Pairwise Label Co-occurrence Prediction (PLCP), and 2) Conditional Label Co-occurrence Prediction (CLCP).
3	Multi-class Text Classification using BERT-based Active Learning	https://arxiv.org/abs/2104.14289			BERT
4	Multi-label classification of promotions in digital leaflets using textual and visual information	https://arxiv.org/abs/2010.03331	Maybe		1) image-based detectionof the descriptions for each individual promotion and 2) multi-label classification of the productcategories using the text from the product descriptions.
	Threshold				
1	Multi-label Classification: A Comparative Study on Threshold Selection Methods	https://www.researchgate.net/publication/277	Yes		
2	Threshold Moving Approaches for Addressing the Class Imbalance Problem and their Application to Multi-label Classification	https://dl.acm.org/doi/abs/10.1145/3441250.3	Maybe		Related to Calibration threshold Th
3	Explainable Automated Coding of Clinical Notes using Hierarchical Label-wise Attention Networks and Label Embedding Initialisation	https://arxiv.org/pdf/2010.15728v4.pdf			
4	Hierarchical multi-label classification using local neural networks	https://www.sciencedirect.com/science/article	Maybe		Hierarchical multi-label classification with local multi-layer perceptron and threshold
1	Multilabel - Text Classification	https://arxiv.org/abs/2003.11644			
2	Multi-Label Text Classification using Attention-based Graph Neural Network	https://arxiv.org/pdf/1905.10070v2.pdf			
3	Label-aware Document Representation via Hybrid Attention for Extreme Multi-Label Text Classification	https://arxiv.org/ftp/arxiv/papers/2010/2010.04			
4	Tag Recommendation for Online Q&A Communities based on BERT Pre-Training Technique	https://arxiv.org/pdf/1905.10070v2.pdf			
5	Label-Wise Document Pre-Training for Multi-Label Text Classification	https://arxiv.org/pdf/1906.02192v1.pdf			
6	Large-Scale Multi-Label Text Classification on EU Legislation	https://arxiv.org/pdf/1906.02192.pdf			
7	Medical Code Prediction from Discharge Summary: Document to Sequence BERT using Sequence Attention	https://arxiv.org/ftp/arxiv/papers/2106/2106.07			
	Hierachical models				
1	TaxoClass: Hierarchical Multi-Label Text Classification Using Only Class Names	https://aclanthology.org/2021.naacl-main.335/		No	Shen et al. (2021)
2	Hierarchy-Aware Global Model for Hierarchical Text Classification	https://aclanthology.org/2020.acl-main.104/		Yes	Zhou et al. (2020)
3	Concept-Based Label Embedding via Dynamic Routing for Hierarchical Text Classification	https://aclanthology.org/2021.acl-long.388.pdf		No	Wang et al. (2021), they cited Zhou et al. (2020)
4	LA-HCN: Label-based Attention for Hierarchical Multi-label TextClassification Neural Network	https://arxiv.org/abs/2009.10938		No	Zhang et al. (2021)
5	Explainable Automated Coding of Clinical Notes using Hierarchical Label-wise Attention Networks and Label Embedding Initialisation	https://arxiv.org/abs/2010.15728		Yes	Dong et al. (2021)
6	Hierarchical Multi-label Text Classification: An Attention-based Recurrent Network Approach	https://dl.acm.org/doi/10.1145/3357384.3357		Yes	Huang et al. (2019)