ICASSP '25 notes and interesting posters

Benedikt Kantz

November 6, 2025

1 Montag Vormittag

1.1 Tutorial: Generative AI and Model Optimization

Problem: (compute) cost, current foundation models not sustainable Solutions:

1.1.1 Sparsity

$\rightarrow$ scalability, less overfitting, interpretability, adaptive ways to introduce sparsity

- post training: optimal brain damage (OBD)/ optimal brain surgery (OBS)
    - dropout by contribution to error, scale by Hessian $\mathcal{H}$ contribution
- training:
    - L1-loss: Convex optim.; no free lunch: initial model very large!, more eqs.
    - exaustive: very expensive
    - greedy/evolutionary solutions: StOMP, GOMP based on L0-norm, but very effective
- pre-training
    - SET
    - randomly initial init $\rightarrow$ evolutionary
- architecutral: grow and shrink networks...

Problem: doesn't really work with LMs (empirical study), but well for other networks (esp. low-weight dropout)

1.1.2 Compression

- filter: storage compresion
- low rank factorization ($\neq$ LoRA), during train time not fine-tuning
- knowledge distillation

2 Dienstag Nachmittag

2.1 Talk: Underwater Communications

- Problem: very slow comm underwater, $\approx$10 kHz range
- Towards moving target, Doppler correction using active SP correction, very manual work

Comment: interesting manual process, tedious work to sample

3 Mittwoch Nachmittag

3.1 Talk: AI+SP

Comment: some basics on diffusion/transformers, a little bit of SP in NNs

4 Donnerstag Vormittag

4.1 Talk: Multiomics

- Genomics: DNA understanding

- Transcriptomics: DNA-¿RNA understanding

- Proteomics: RNA-¿Protein structures

- Knowledge graphs: how do these systems influnce each other

- Flow:

  - identify DNA mutation that triggers illness
  - find possible RNA mechanism
  - find good fitting small ring structure
  - $\rightarrow$ check for side effects in knowledge graph! (certain protein effects unwantend)
  - $\rightarrow$ then test $\rightarrow$ animal tests, reduce through ML!

- Graph diffusion for drug discovery: noise schedule for diffusion essential, i.e. cosine-square schedule

  - diffuse graphs from atoms & edges as adjacency matrix
  - what is noise: discrete noise: each atom is discrete state $\implies$ graph structure undergoes state transition change
  - naive: uniform structure, not really chemically sensible - conditional probabibilites $\implies$ not uniform but marginal distribution of molecules in training (just logical!), same for edge (with deletion!)
  - one step further: consider carbon rings, restriction based on maximum bonds of atom (freie radikale)
  - SMILE-file, QED: Quantitative Esitmate of Drug likeness (from RDKit)
  - Existing methods: Time-consuming, progress slow, very few good molecules
  - Their work: jointly perturb rings+nodes
  - other approaches: motives as super-node with rings, difficulty: ring attachments - only ≈ 1 % improvement!
  - novelty however high, one molecule of them even patented!

- Knowledge graphs:

  - GNN link prediction
  - none of the existing benchmarks include features!
  - maybe talk to author!

Comment: focused on drug discovery using diffusion, not much on multiomics. . .

5 Lectures/Orals

Table 1:

Lecture	URL	Notes
Diversity-Seeking Techniques for Red-Teaming LLMs	`https://ieeexplore.ieee.org/document/10890844/`	Add RL-Loss to train similar to GAN by backpropagating if the model returns very similar output (i.e. discrimimnator)-¿ Very fragile learning; Limited further studies
FDR Control for Complex-Valued Data	`https://ieeexplore.ieee.org/document/10889705`	similar to LASSO; sparsifying system under certain guarantees
SpectralCam: High-Resolution Low-Cost Spectral Imaging Using DSLR Cameras	`https://ieeexplore.ieee.org/document/10887725`	Interesting concept of applying photo filter to DSLR sensor, Bayesian pattern restoration "learned" using diffusion & attenuation mtx
Fusing Multimodality of Large Language Models and Satellite Imagery	`https://ieeexplore.ieee.org/document/10889624`	Could be interesting in combination with HEREDITARY geospatial data once we have access
Controllable Forgetting Mechanism for Few-Shot Class-Incremental Learning	`https://arxiv.org/pdf/2501.15998`	Using embedding space to classify, add new classifier based on distance, seems rather hyperparameter-sensitive

6   Posters

— **Poster** — **Title** — **Information** — — — — — — — — —![poster]($out_reduced/IMG_2777.png$)|$PD-$
$VOST: Parkinson's Disease Voice Spectrogram Transformer|*lliasTougui, MehdiZakroum, OuassimKarakchou, Mounir$
$Parkinson's Disease Voice Spectrogram Transformer](https://ieeexplore.ieee.org/abstract/document/10889820/)| |![p$
$pee/ginbergaepf.ch, facia, surya, gauraviaroattydefenderal*, [TheEPFLcombinationalbenchmarksuite](https:$
$//si2.epfl.ch/demichel/publications/archive/2015/IWLS15.pdf)| |![poster](out_reduced/IMG_2749.png)|Latentdiffusio$
$MicheleMancusi, YuriiHalychanskyi, KinWaiCheuk, EloiMoliner, Chieh-HsinLai, StefanUhlich, JunghyunKoo.*, [Laten$
$//ieeexplore.ieee.org/abstract/document/10890708/)| |![poster](out_reduced/IMG_2761.png)|doacrowendusportedtenba$
$doacrowendusportedtenbandmentorineatpearenoses.Thepipeline*, [doacrowendusportedtenbandmentorineatpearenose$
$outperformingmodelslikeCV-SOG, Motifs, andVCTree.*, [EFFICETHREPARAMNE]()| |![poster](out_reduced/IMG_2$
$Explainablemultimodalmusicunderstanding|*tasks, achieving57.34\%forgenreand48.53\%foremotionClassi-*, [Musiclin$
$Explainablemultimodalmusicunderstanding](https://ieeexplore.ieee.org/abstract/document/10889771/)| |![poster](o$
$IC: On-the-SphereLearnedOmnidirectionalImageCompressionwithAttentionModulesandSpatialContext|*$
$Bidgoll, PascalFrossard?, AndréKaup", ThomasMaugey?*, [OSLO-IC: On-the-SphereLearnedOmnidirectionalImage$
$//ieeexplore.ieee.org/abstract/document/10889131/)| |![poster](out_reduced/IMG_2649.png)|Exploitingtherelationshipw$
$QiuboMa', HangYu\%, YuanShan3, PinzhuoTian1*, [Exploitingtherelationshipwithintheunlabelledsamplesbysetmatching$
$//ieeexplore.ieee.org/abstract/document/10889522/)| |![poster](out_reduced/IMG_2662.png)|EvaluatingContrastiveMe$
$methods[1, 2]andnovelhybridapproaches*, [EvaluatingContrastiveMethodologiesforMusicRepresentationLearningUsi$
$//ieeexplore.ieee.org/abstract/document/10888157/)| |![poster](out_reduced/IMG_2666.png)|Fine-tuningandpromptopt$
$Twogreatstepsthatworkbettertogether|*DongSun, WenyaGuo, XumengLiu, YingZhang*, ZhaoxiangHou, ZengxiangLi*, [$
$Twogreatstepsthatworkbettertogether](https://arxiv.org/abs/2407.10930)| |![poster](out_reduced/IMG_2667.png)|Diag$
$DrivenBearing-FaultDetectioninInductionMotorandDrivesusingGraphSamplingandAggregationNetwork|*$
$HaraprasadBadajena, SuryanarayanMajhi, BivashChakraborty, MamataJenamani, AurobindaRoutray, RonitDutta*, [$
$//ieeexplore.ieee.org/abstract/document/10889484/)| |![poster](out_reduced/IMG_2665.png)|HybridFeatureFusionforb$
$YiZhu', XiangyangLiu!?, TianqiPang', XuncanXiao!, XiaofanZhang33, ChenyouFan!.**, [HybridFeatureFusionforEnha$
$//ieeexplore.ieee.org/abstract/document/10888956/)| |![poster](out_reduced/IMG_2664.png)|Texttomusicaudiogeneratio$
$Are-engineeringofaudioldmmodel|*1nhtonneWegnerNetealPonesin, DerenHertemans, RoggerWattenhofer*, [Texttomu$
$Are-engineeringofaudioldmmodel](https://www.diva-portal.org/smash/record.jsf?pid=diva2:$
$1845150)| |![poster](out_reduced/IMG_2655.png)|ExploringtheDistributionofCellSubpopulationsinPancreaticDuctalAd$
$YagiDeng, WenjieCai, BentaoSong, BinYang, LingmingKong, QingfengWang \gg, JunHuang*, [ExploringtheDistributiona$
$//ieeexplore.ieee.org/abstract/document/10890640/)| |![poster](out_reduced/IMG_2641.png)|ClassificationofEye-$
$TrackingDataBasedonSpatiotemporalAttentionEncoding|*MajuHel, ChenXia't, KuanL, TanZhangt*, [Classificationof$
$//ieeexplore.ieee.org/abstract/document/10889246/)| |![poster](out_reduced/IMG_2669.png)|GPT-LAD:$
$LeveragingLargeMultimodalModelsforLogicalAnomalyDetection|*MainPerformanceonMVTecLOCO(A:$
$all, L: logical, S: structural)*, [GPT-LAD: LeveragingLargeMultimodalModelsforLogicalAnomalyDetection](https$
$//ieeexplore.ieee.org/abstract/document/10888757/)| |![poster](out_reduced/IMG_2734.png)|Virtualmultipleximmunohi$
$applicationoncellblockofeffusionandaspirationcytology|*XianchaoGuan12, ZhengZhang?, YifengWang?, YeZhang?, L$
$applicationoncellblockofeffusionandaspirationcytology](https://onlinelibrary.wiley.com/doi/abs/10.1002/dc.2434$
$ViesinanAudeSpectresunTaedome981ontheSpeechcommanda*, [ViesinanAudeSpectresunTaedome981ontheSpeechcomm$
$YagiDeng, McnjeCai, BenaoSong, BinYang, LimphingKung, QedengPung \gg, JamH*, [ExploringtheDistributionofCellSu$
$//ieeexplore.ieee.org/abstract/document/10890640/)| |![poster](out_reduced/IMG_2654.png)|NanoGen:$
$AHigh-affinityNanobodyGenerationModelwithGuidedDiffusion|*DezhiWu', XuejiaoLiu?, YimingQin", StephanieM.L$
**, [NanoGen: AHigh-affinityNanobodyGenerationModelwithGuidedDiffusion](https://ieeexplore.ieee.org/abstra$
$Acomprehensiveknowledgebaseofmanuallycuratedpeptidesagainstdenguevirus|*Emas: 2230112006.M.00u.59.congna00$
$Acomprehensiveknowledgebaseofmanuallycuratedpeptidesagainstdenguevirus](https://link.springer.com/article/10$
$YifeiChen, ShenghaoZhu.ZhaojieFang.ChangLiu, BinfengZou, LinweiQiu, YuheWang.ShuoChang.FanJia, FelwelQin" .$
$//ieeexplore.ieee.org/abstract/document/10888363/)| |![poster](out_reduced/IMG_2647.png)|ExperimentsdemonstratoM$
$MACachieves22|![poster](out_reduced/IMG_2653.png)|Robustkernelsparsesubspaceclustering|*dacesta54, P.O.Box180,$
$//ieeexplore.ieee.org/abstract/document/10888170/)| |![poster](out_reduced/IMG_2733.png)|Agenerative-$
$augmenteddeepmatrixfactorizationmodelforpoirecommendations|*ChonseLio, HonglZhange, AestaMan, MasaZhang*,$
$//ieeexplore.ieee.org/abstract/document/10890082/)| |![poster](out_reduced/IMG_2732.png)|2011IEEEInternationalCo$
$KunMa12, QilongHan', JingzhengYao?.*, ChangmaoWu'andChunrulNazA*, [2011IEEEInternationalConferenceonAcou$
$//ieeexplore.ieee.org/abstract/document/5947741/)| |![poster](out_reduced/IMG_2652.png)|Maskedimagemodelingadv$
$Derformanceandpromisingscalabilityinmedical MIM.Spectically, weproposeatelliterativemaskingstratcoybasedon*, [$
$//openaccess.thecvf.com/content/WACV2023/html/Chen\_Masked\_Image\_Modeling\_Advances\_3D\_Medical\_Image\_And$
$Losses.....*Inlt..*, [Latentwatermarkingofaudiogenerativemodels](https://ieeexplore.ieee.org/abstract/document/10$
$Hybridmemoryandattentionnetworkforunsuperviseddanomalydetection|*JiahaoLi, YigiangChen, YunbingXing, YangGu, X$

Hybridmemoryandattentionnetworkforunsupervisedanomalydetection](*https* : *//ieeexplore.ieee.org/abstract/document*

*Asshowninthe*29, 3*P*%, *and*4∗, [*Performanceevaluationofobjectdetectionandtrackingsystems*](*https* : *//www.academia*

15)| |![*poster*](*out$_r$educed/IMG$_2$*731.png*)|*XD*0*h*18122089*rojectGenderEqualtyMonitor* − *ANIR* − 19 −

*GE*30−0012).*Itwasportormed*|∗*FrançoisBuot′*, *CamilleGuinaudeau*?, *CyrilGrouin*?, *SaharGhannay*?, *Shin′ichiSatoh* ≫

∗, [*XD*0*h*18122089*rojectGenderEqualtyMonitor*−*ANIR*−19−*GE*30−0012).*Itwasportormed*]()| |![*poster*](*out$_r$educed/IM*

13)*Olfonses, Alexandre, JadeCopet, GabrielSynarve, andYestAdt.″Highfidelity*∗, [*Latentwatermarkingofaudiogenerat*

*//ieeexplore.ieee.org/abstract/document/*10889782/)| |![*poster*](*out$_r$educed/IMG$_2$*651.png*)|*UnsupervisedDomainAdap*

TWDR.&methodforremovingirrelevantdatafromatrainingset∗, [*UnsupervisedDomainAdaptationViaDataPruning*](*http*

*//ieeexplore.ieee.org/abstract/document/*10890190/)| |![*poster*](*out$_r$educed/IMG$_2$*756.png*)|*Investigationofwhisperas*

speechaudio|∗*AGHUniversityofKrakow.Poland*∗, [*Investigationofwhisperasrhallucinationsinducedbynon*−*speechaud*

*//ieeexplore.ieee.org/abstract/document/*10890105/)| |![*poster*](*out$_r$educed/IMG$_2$*757.png*)|*Retrievalaugmentedcorre*

ErnestPusateri, AnmolWalia, AniruchKashi, BortikBandyopadhyay, NadiaHyder, SayantanMahinder, RavitejaAnantha

∗∗, [*Retrievalaugmentedcorrectionofnamedentityspeechrecognitionerrors*](*https* : *//ieeexplore.ieee.org/abstract/docu*

*PromptingDrivenSAM*2*for*3*DMedicalImageSegmentation*|∗*aSortedindexfromC* : *s*, −(6*r*, 62..., 62..., 4)∗, [*Self*−*Prompti*

*//ieeexplore.ieee.org/abstract/document/*10889344/)| |![*poster*](*out$_r$educed/IMG$_2$*778.png*)|*IntegratingPauseInforma*

SpeechandAudioTechnologylab, TsinghuaUniversity, China∗, [*IntegratingPauseInformationwithWordEmbeddingsinLa*

*//ieeexplore.ieee.org/abstract/document/*10888563/)| |![*poster*](*out$_r$educed/IMG$_2$*750.png*)|*Multi*−*scaleContextIntertv*

*YeZhang′*2, *XlanchaoGuan*13, *HengrulLI″, XiangmingYan″, ZiyueWang″, YongbingZhang′*∗, [*Multi*−*scaleContextIntertv*

*//ieeexplore.ieee.org/abstract/document/*10889659/)| |![*poster*](*out$_r$educed/IMG$_2$*779.png*)|*SELMA* :

ASpeech−*EnabledLanguageModelforVirtualAssistantInteractions*|∗*SimplifiedPigstina, Meducescomplarty*∗, [*SELMA*

ASpeech−*EnabledLanguageModelforVirtualAssistantInteractions*](*https* : *//ieeexplore.ieee.org/abstract/document/1*

neuronalinteractionsinAlzheimer′sdisease : *thepotentialroleofa′cytokinecycle′indiseaseprogression*| ∗

*hutcmoshini*001@*Bentuedusg, choc*0010@*e.ntu.edusg, yihao*001@*e.ntu.edusg, conghao*001@*e.ntu.edusg, asjagath@ntu.ed*

neuronalinteractionsinAlzheimer′sdisease : *thepotentialroleofa′cytokinecycle′indiseaseprogression*](*https* :

*//onlinelibrary.wiley.com/doi/abs/*10.1111/*j*.1750−3639.1998.*tb*00136.*x*)| |![*poster*](*out$_r$educed/IMG$_2$*753.png*)|*Wireles*

13*thChinaConference, CWSN*2019, *Chongqing, China, October*12–14, 2019, *RevisedSelectedPapers*|∗1.*CollegeofCompu*

13*thChinaConference, CWSN*2019, *Chongqing, China, October*12–14, 2019, *RevisedSelectedPapers*](*https* :

//books.google.com/books?*hl* = *en&lr* = &*id* = *o*3*fADwAAQBAJ&oi* = *fnd&pg* = *PR*6&*dq* = 1. +

College + *of* + *Computer* + *Science* + *and* + *Technology*, +*Chongging* + *University* + *of* + *Posts* + *and* +

Telecommunications, +*Chongging*.+*China&ots* = *xK*5*tdJxqE*6&*sig* = 49*RR*4*LsIhUDMwMlUVyyZO*$_-$*Et*08*I*)| |![*poster*](*o*

GeneratedMusicDetectionanditsChallenges|∗*Suno, Udio, Riffusion, ...*∗, [*AI*−*GeneratedMusicDetectionanditsChalleng*

*//ieeexplore.ieee.org/abstract/document/*10890655/)| |![*poster*](*out$_r$educed/IMG$_2$*785.png*)|*ATraining*−

FreeCorrelation−*WeightedModelforZero*−/*Few*−*ShotIndustrialAnomalyDetectionwithRetrievalAugmentation*|∗

WeiRan, ZefangYu, SunchengXiang, TingLiu, YuzhuoFu∗, [*ATraining*−*FreeCorrelation*−*WeightedModelforZero*−/*Few*

*//ieeexplore.ieee.org/abstract/document/*10890083/)| |![*poster*](*out$_r$educed/IMG$_2$*746.png*)|*Parameter*−

efficientfine−*tuningoflarge*−*scalepre*−*trainedlanguagemodels*|∗*ChangzengFits, ZelinFut, ShaojunYant.XiaoyongLyv*

*//www.nature.com/articles/s*42256−023−00626−4)|