Course Syllabus

Tentative Schedule: Minor changes may be made as needed throughout the quarter.

Lecture #	Title	Readings
wk1: Lecture 1: October 1 Course Intro	Course overview & logistics (https://canvas.ucsd.edu/courses/58846/files/13151386?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13151386/download?download_frd=1)	• C. Manning: Linguistics & Deep Learning (https://aclanthology.org/J1 • PyTorch Tutorial from Stanford CS224N (https://colab.research.google.com/drive/1AiBuEL_V9u16jGMctpYN3Byx76_usp=sharing)
wk1: Lecture 2: October 3 Intro to Neural NLP	Machine Learning foundations: linear models (https://canvas.ucsd.edu/courses/58846/files/13164617?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13164617/download?download_frd=1)	• **[Eisenstein]** 2.1; 2.5.0 - 2.5.1 • A Few Useful Things to Know About Machine Learning (https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf)
wk2: Lecture 3: October 8 Intro to Neural NLP	Neural Networks foundations: feedforward neural networks (https://canvas.ucsd.edu/courses/58846/files/13206214?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13206214/download?download_frd=1)	• **[Eisenstein]** 3.0-3.3 • Deep Averaging Networks (https://aclanthology.org/P15-1162.pdf) • Natural Language Processing with Small Feed-Forward Networks (https://aclanthology.org/D17-1309.pdf)
wk2: Lecture 4: October 10 Intro to Neural NLP	Word Embeddings & Tokenization (https://canvas.ucsd.edu/courses/58846/files/13226847?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13226847/download?download_frd=1)	• Mikolov et al. 2013 word2vec (https://arxiv.org/pdf/1301.3781.pdf) • Pennington et al. 2014 GloVe (https://nlp.stanford.edu/pubs/glove.pdf • Tokenization: Sennrich et al. 2016, Byte Pair Encoding (https://aclar 1162.pdf)
wk3: Lecture 5: October 15 Modern LMs: Background	N-gram & Neural Language Models (RNNs) (https://canvas.ucsd.edu/courses/58846/files/13283440?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13283440/download?download_frd=1)	Language Modeling, Collins notes (https://canvas.ucsd.edu/courses/58846/ wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13270746/download?d Bengio et al. 2003 (https://www.jmlr.org/papers/volume3/bengio03a/bengio03
wk3: Lecture 6: October 17 Modern LMs: Background	Neural Machine Translation & Cross-Attention in Seq2Seq (https://canvas.ucsd.edu/courses/58846/files/13300495?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13300495/download?download_frd=1)	Sutskever et al. 2024 (https://proceedings.neurips.cc/paper_files/paper/2014/file/a14ac55a4f27472c5d Paper.pdf) Other readings: Neubig 2017 (https://arxiv.org/pdf/1703.01619)

wk4: Lecture 7: October 22 Modern LMs: Key ingredients	**Self-Attention & Transformers** (https://canvas.ucsd.edu/courses/58846/files/13347732?wrap=1) ↓ (https://canvas.ucsd.edu/courses/58846/files/13347732/download?download_frd=1) The notebook shown in class consists of snippets from the tutorial, **refer to the full tutorial instead** (https://www.youtube.com/watch?v=kCc8FmEb1nY&t=17s&ab_channel=AndrejKarpathy)	**Vaswani et al. 2017** (https://arxiv.org/pdf/1706.03762.pdf) **Karpathy 2023, Let's build GPT** (https://www.youtube.com/watch?v=kCc8FmEb1nY&ab_channel=AndrejKarpathy) **Illustrated Transformer** (https://jalammar.github.io/illustrated-transformer/) **Beltagy et al. 2020** (https://arxiv.org/pdf/2004.05150.pdf)
wk4: Lecture 8: October 24 Modern LMs: Key ingredients	**Pretraining: Encoders (BERT/ELECTRA/DeBERTa)** (https://canvas.ucsd.edu/courses/58846/files/13374001?wrap=1) ↓ (https://canvas.ucsd.edu/courses/58846/files/13374001/download?download_frd=1) + **Project information**	**Peters et al. 2018 ELMo** (https://arxiv.org/pdf/1802.05365.pdf) **Devlin et al 2019 BERT** (https://arxiv.org/abs/1810.04805) **Clark et al 2020 ELECTRA** (https://arxiv.org/pdf/2003.10555.pdf) **GHe 2021 DeBERTa** (https://arxiv.org/pdf/2006.03654.pdf)
wk5: Lecture 9: October 29 Modern LMs: Key ingredients	**Pretraining part 2: Decoders (GPT); Decoding Methods; Scaling Laws** (https://canvas.ucsd.edu/courses/58846/files/13425514?wrap=1) ↓ (https://canvas.ucsd.edu/courses/58846/files/13425514/download?download_frd=1)	**GPT2** (https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf) **Holtzman 2019 Nucleus Sampling** (https://arxiv.org/pdf/1904.09751.pdf)
wk5: Lecture 10: October 31 Modern LMs: Key ingredients	**Prompting ; Instruction Tuning** (https://canvas.ucsd.edu/courses/58846/files/13447794?wrap=1) ↓ (https://canvas.ucsd.edu/courses/58846/files/13447794/download?download_frd=1)	**Brown et al. 2020 GPT3** (https://arxiv.org/pdf/2005.14165.pdf) **Min et al. 2022 Rethinking Demonstrations** (https://arxiv.org/pdf/2202.12 **Sanh et al 2022 Instruction Tuning** (https://arxiv.org/pdf/2110.08207) **FLAN-T5** (http://Scaling%20Instruction-Finetuned%20Language%20Models)
wk6: Lecture 11: November 5 Modern LMs: In practice	**Retrieval Augmented Generation (RAG)** (https://canvas.ucsd.edu/courses/58846/files/13497079?wrap=1) ↓ (https://canvas.ucsd.edu/courses/58846/files/13497079/download?download_frd=1) **Project Proposal Template (due but optional)** (https://www.overleaf.com/read/mqjvhgywmykq#c9f88a)	**Lewis et al. 2020 RAG** (https://proceedings.neurips.cc/paper/2020/file/6b493230205f780e1bc26945df74 **Karpukhin et al. 2020 DPR** (https://aclanthology.org/2020.emnlp-main.550.pd **Borgeaud et al. 2022 RETRO** (https://arxiv.org/abs/2112.04426)
wk6: Lecture 12: November 7 Modern LMs: In practice	**Parameter Efficient Fine-Tuning (PEFT) (** (https://canvas.ucsd.edu/courses/58846/files/13519902?wrap=1) ↓ (https://canvas.ucsd.edu/courses/58846/files/13519902/download?download_frd=1) Lottery Tickets/LoRA/Adapters)	**Frankle & Cabin 2019, Lottery Ticket Hypothesis** (https://arxiv.org/pdf/1803 **Ansell et al 2022, Sparse-finetuning** (https://aclanthology.org/2022.acl-long. **Hu et al 2021, LoRA** (https://arxiv.org/pdf/2106.09685) **Houlsby et al 2019, Adapter functions** (https://proceedings.mlr.press/v97/houlsby19a/houlsby19a.pdf) **He et al 2022** (https://openreview.net/pdf?id=0RDcd5Axok)

wk7: Lecture 13: November 12 Modern LMs: In practice	**Knowledge Representation in Transformer LLMs (parametric memory)** (https://canvas.ucsd.edu/courses/58846/files/13562215?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13562215/download?download_frd=1)	(https://aclanthology.org/2021.emnlp-main.446.pdf) Geva et al. 2021 FeedForward Memory (https://aclanthology.org/2021.emnlp Meng et al. 2022 ROME (https://proceedings.neurips.cc/paper_files/paper/2022/file/6f1d43d5a82a37e89b Paper-Conference.pdf)
wk7: Lecture 14: November 14 Modern LMs: In practice	**Code Generation with LLMs** (https://canvas.ucsd.edu/courses/58846/files/13590911?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13590911/download?download_frd=1)	Chen et al 2021 Codex (https://arxiv.org/pdf/2107.03374) Li et al 2022 AlphaCode (https://arxiv.org/pdf/2203.07814) Ahn et al 2022 SayCan (https://say-can.github.io/assets/palm_saycan.pdf)
wk8: Lecture 15: November 19 Modern LMs: In practice	**Interpretability of Neural NLP** (https://canvas.ucsd.edu/courses/58846/files/13643477?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13643477/download?download_frd=1) (Probes/Sparse Autoencoders/Dataset Artifacts)	Huben et al, ICLR 2024 (https://openreview.net/forum?id=F76bwRSLeK) Levy et al, ACL 2024 (https://aclanthology.org/2024.acl-long.818.pdf) Conneau et al, 2018 (https://aclanthology.org/P18-1198.pdf) (https://aclanthology.org/2021.emnlp-main.446.pdf) Belinkov 2020 (https://aclanthology.org/2022.cl-1.7.pdf)
wk8: Lecture 16: November 21 Modern LMs: In practice	**Question Answering** (https://canvas.ucsd.edu/courses/58846/files/13671319?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13671319/download?download_frd=1)	Chen et al 2017 DrQA (https://arxiv.org/pdf/1704.00051) Seo et al 2019 BiDAF (https://arxiv.org/pdf/1611.01603)
wk9: Lecture 17: November 26 Modern LMs: In practice	**LLMs & Society** (https://canvas.ucsd.edu/courses/58846/files/13719536?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13719536/download?download_frd=1) • Privacy & LLMs • Bias/Incivility/Unethical use • Safeguarding LLMs	Carlini et al, 2021: extracting data from LLMs (https://www.usenix.org/system/files/sec21-carlini-extracting.pdf) Jurgens et al 2018 (https://aclanthology.org/P19-1357.pdf) Bender et al 2021, stochastic parrots (https://dl.acm.org/doi/10.1145/344218
wk9: Lecture 18: November 28	Thanksgiving - no lecture	
wk10: Lecture 19: December 3	**Slides** (https://canvas.ucsd.edu/courses/58846/files/13780378?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13780378/download?download_frd=1) • LLMs practical considerations	

Modern LMs: In practice	• Open Problems in NLP • Beyond English • Course Wrap-Up	
wk10: Lecture 20: December 5 Modern LMs: In practice	No lecture: work on projects, or AMA in CSE 4109	
December 06 PROJECT	**Project Proposal Template** (https://www.overleaf.com/read/mqjvhgywmykq#c9f88a) **Final Project Report Template** (https://www.overleaf.com/read/mdkmcdjmmzgh#030872)	

Readings. The assigned readings are intended to complement the lectures. You may consult them before the lecture to prepare or after the lecture to review the material. The majority of our readings will be technical papers, as we do not have a designated textbook. However, for those interested in additional references, the following books are recommended:

Book 1: **[Goldberg]** Goldberg: A Primer on Neural Network Models for Natural Language Processing (https://u.cs.biu.ac.il/~yogo/nnlp.pdf) .

Book 2: **[Eisenstein]** Eisenstein: Natural Language Processing (https://canvas.ucsd.edu/courses/58846/files/13052134?wrap=1) (https://canvas.ucsd.edu/courses/58846/files/13052134/download?download_frd=1)

Book 3: **[ZLLS]** Aston Zhang, Zack C. Lipton, Mu Li, Alex J. Smola, Dive into Deep Learning (https://d2l.ai/index.html)

Book 4: **[J&M]** Jurafsky and Martin: Speech and Language Processing (3rd ed. draft) (https://web.stanford.edu/~jurafsky/slp3/)

Course Summary:

Date	Details	Due
Fri Oct 11, 2024	The #FinAid Quizz (https://canvas.ucsd.edu/courses/58846/assignments/864068)	due by 11:59pm
Sun Oct 20, 2024	PA1 (https://canvas.ucsd.edu/courses/58846/assignments/864875)	due by 11:59pm

Date	Details	Due
Wed Oct 30, 2024	Quiz 1: Lecture 2-7 (https://canvas.ucsd.edu/courses/58846/assignments/869682)	due by 11:59pm
Wed Nov 6, 2024	PA2 (https://canvas.ucsd.edu/courses/58846/assignments/868469)	due by 11:59pm
Mon Nov 18, 2024	PA3 (https://canvas.ucsd.edu/courses/58846/assignments/872209)	due by 11:59pm
Thu Nov 21, 2024	Quiz 2: up to Lecture 14 (https://canvas.ucsd.edu/courses/58846/assignments/875615)	due by 11:59pm
Thu Nov 28, 2024	Quiz 3: up to Lecture 16 (https://canvas.ucsd.edu/courses/58846/assignments/875623)	due by 11:59pm
Mon Dec 2, 2024	PA4 (https://canvas.ucsd.edu/courses/58846/assignments/876197)	due by 11:59pm
Mon Dec 9, 2024	Quiz #4 (optional - in case you forgot quiz 1/2/3) (https://canvas.ucsd.edu/courses/58846/assignments/879615)	due by 11:59pm