

コース名	AI・データサイエンスコース		
講座名	—		
科目名	LLM（大規模言語モデル）/生成AIのための深層学習 （トランスフォーマー編）		
必修・選択	—	単位	—
概要・目的	<p>トランスフォーマーや拡散モデルは、CNNやRNNと同じ深層学習の一学習手法として示されるが、これらの原理は大きく異なる。CNNやRNNを第一世代の深層学習と呼ぶならば、トランスフォーマーや拡散モデルは、世代が完全に変わった第二世代の深層学習と呼ぶにふさわしい。</p> <p>どう違うかは、講義内でゆっくり説明するが、ChatGPTなどを思い浮かべてもらえれば、その一端は推測できる。ChatGPTは、</p> <p>（1）超大規模なラベルなしのデータで学習しているという部分が一つ目の違いで、（2）これで解けるタスクは一つでなく複数で幅広いというところが二つ目の違いとなる。</p> <p>これと比較して、第一世代の深層学習は、ラベル付きのデータで学習して、タスクは一つが通常である。あまり、この違いを強調して教える人もいないが、ここではここを中心軸として教えたい。</p>		
到達目標	ChatGPTの背景にある深層学習技術を理解すること、ラベルなしで学習するトランスフォーマーの事前学習モデルを理解すること、トランスフォーマーの事前学習モデルを用いた下流タスクを理解すること		
授業方法	講義	実施形態	遠隔同期
評価方法	数回行う小テストによる		
授業項目	1	LLM/トランスフォーマーのイントロ：ChatGPTとは何か、生成モデルとは何か、フィードフォワード層	
	2	LLM/トランスフォーマーの学習：活性化関数、トークン化、正規化層、スキップ接続	
	3	LLM/トランスフォーマーの学習：自己注意機構、エンコーダーとデコーダー	
	4	LLM/トランスフォーマーの学習：下流タスクの学習	
	5	LLM/トランスフォーマーの学習：ファインチューニング、インコンテキスト学習	
	6	LLM/トランスフォーマーの学習：事前学習モデルの学習の目的（次単語予測、マスク言語モデル、対比学習）	
	7	ビジョントランスフォーマーの学習：プレテキストタスク、マスク画像モデル	
	8	LLM/トランスフォーマーを取り巻く諸々の問題：自己教師あり学習的な視点、線形のトランスフォーマー、実行環境、LangChain, copilotなど	
	9		
	10		
	11		
	12		
	13		
	14		
	15		
使用教材	講師がスライドなどで提供する		
特記事項			