ขอบเขตการสอบกลางภาค — สำหรับนักศึกษา

NATURAL LANGUAGE PROCESSING WITH DEEP LEARNING (2026)

คณะเทคโนโลยีสารสนเทศ สจล. · ผู้สอน: ผศ. ดร. นนท์ คนึงสุขเกษม

รูปแบบข้อสอบ

หัวข้อ	รายละเอียด
จำนวนข้อ	80 ข้อ · ข้อละ 0.5 คะแนน
สัดส่วนคะแนน	40% ของคะแนนทั้งรายวิชา
เวลา	3 ชั่วโมง
ตอนที่ 1	ปรนัย 65 ข้อ (5 ตัวเลือก A–E เลือกคำตอบที่ถูกที่สุด 1 ข้อ)
ตอนที่ 2	เขียนตอบ / เติมคำ 15 ข้อ (ต้องแสดงวิธีคิดเมื่อเป็นข้อคำนวณ)
เนื้อหา	สัปดาห์ที่ 1–7

สิ่งที่อนุญาตให้นำเข้าห้องสอบ

	รายการ
✅	กระดาษ A4 **1 แผ่น** ที่ **เขียนด้วยลายมือตัวเองเท่านั้น** (ใช้ได้ทั้งสองหน้า)
✅	**เครื่องคิดเลข**
❌	โทรศัพท์มือถือ · โน้ตบุ๊ก · แท็บเล็ต · อุปกรณ์เชื่อมต่ออินเทอร์เน็ตทุกชนิด

ภาพรวม: จำนวนข้อแยกตามสัปดาห์

สัปดาห์	หัวข้อหลัก	จำนวนข้อ	สัดส่วน
1	Intro & Word Representation	11	13.75%
2	RNN / LSTM / GRU	12	15.00%
3	Seq2Seq & Attention	11	13.75%
4	Transformer Encoder	12	15.00%
5	Transformer Decoder	11	13.75%
6	Encoder — BERT	12	15.00%
7	Decoder — GPT	11	13.75%

สัปดาห์	หัวข้อหลัก	จำนวนข้อ	สัดส่วน
	รวม	**80**	**100%**

> ทุกสัปดาห์มีน้ำหนักใกล้เคียงกัน (11–12 ข้อ) **ไม่มีสัปดาห์ใดที่ข้ามได้** หัวข้อที่กำกับ *(ข้อรวม)* คือหัวข้อที่ปรากฏเป็นตัวเลือกภายในข้อแบบ "ข้อใดต่อไปนี้ถูกต้อง" — ต้องอ่านเช่นกัน

📘 สัปดาห์ 1 — Intro & Word Representation · รวม 11 ข้อ

#	หัวข้อย่อย	ข้อ
1.1	การแทนความหมายของคำ · การลดมิติ (One-Hot, WordNet, SVD, LSA) *(ข้อรวม)*	1
1.2	Co-occurrence Matrix · PMI · PPMI	2
1.3	Word2Vec (Skip-gram · CBOW · Objective Function)	2
1.4	Subword Tokenization (BPE · WordPiece · Unigram)	3
1.5	SentencePiece (และการใช้กับภาษาไทย)	2
1.6	PyTorch: Loss Functions (NLLLoss / CrossEntropyLoss)	1

📗 สัปดาห์ 2 — RNN / LSTM / GRU · รวม 12 ข้อ

#	หัวข้อย่อย	ข้อ
2.1	ข้อจำกัดของโมเดลก่อน RNN (N-gram · Fixed-Window NLM) · GRU *(ข้อรวม)*	1
2.2	RNN & Parameter Sharing	1
2.3	Weight Initialization (Xavier / He)	1
2.4	สถาปัตยกรรมตามลักษณะงาน (Many-to-One / Many-to-Many)	1
2.5	Perplexity	1
2.6	BPTT · Vanishing / Exploding Gradient · Gradient Clipping	2
2.7	LSTM (Cell State · Forget / Input / Output Gate)	3
2.8	Bidirectional RNN	1
2.9	Teacher Forcing & Exposure Bias	1

📙 สัปดาห์ 3 — Seq2Seq & Attention · รวม 11 ข้อ

#	หัวข้อย่อย	ข้อ
3.1	BLEU (Clipped Precision · Brevity Penalty)	2
3.2	ROUGE · ROUGE-L (LCS)	1

#	หัวข้อย่อย	ข้อ
3.3	Decoding: Greedy · Beam Search · Length Normalization	1
3.4	Temperature Sampling	1
3.5	Teacher Forcing Ratio (Scheduled Sampling)	1
3.6	Information Bottleneck → ทำไมต้องมี Attention	1
3.7	คณิตศาสตร์ของ Attention (Context Vector · Softmax)	2
3.8	Query / Key / Value ใน Seq2Seq Attention	1
3.9	Bahdanau (Additive) vs Luong (Dot-Product)	1

📕 สัปดาห์ 4 — Transformer Encoder · รวม 12 ข้อ

#	หัวข้อย่อย	ข้อ
4.1	ทำไมต้องเลิกใช้ RNN / CNN (Sequential Trap · Receptive Field)	1
4.2	Query / Key / Value และ Self-Attention	1
4.3	Scaled Dot-Product (ทำไมต้องหาร $\sqrt{d_k}$)	1
4.4	Positional Encoding (สูตร sin / cos)	2
4.5	Multi-Head Attention · การซ้อนชั้น · W^O	2
4.6	Residual Connection	1
4.7	Layer Normalization (เทียบ Batch Normalization)	1
4.8	Feed-Forward Network · Shape Invariance · การนับพารามิเตอร์	2
4.9	ความซับซ้อนหน่วยความจำ $O(T^2)$	1

📓 สัปดาห์ 5 — Transformer Decoder · รวม 11 ข้อ

#	หัวข้อย่อย	ข้อ
5.1	Causal / Look-ahead Mask	1
5.2	ความต่างระหว่าง Training กับ Inference	1
5.3	Cross-Attention	1
5.4	**KV Cache** (หลักการ · คอขวด · GEMM/GEMV · การคำนวณขนาด) ⭐	**5**
5.5	การวางแผน VRAM	1
5.6	Decoding & Sampling (Top-k · Top-p / Nucleus · ลำดับการกรอง logits)	2

📔 สัปดาห์ 6 — Encoder: BERT · รวม 12 ข้อ

#	หัวข้อย่อย	ข้อ
6.1	Self-Supervised Learning & Transfer Learning (Pretext Task)	1
6.2	Input Embedding (Token + Segment + Position)	1
6.3	`[CLS]` · `[SEP]` · Pooled Output	1
6.4	Masked Language Modeling (MLM)	1
6.5	Next Sentence Prediction (NSP)	1
6.6	Fine-Tuning vs Feature Extraction (BERT vs ELMo)	1
6.7	RoBERTa	1
6.8	Extractive QA / Span Selection (SQuAD)	1
6.9	DeBERTa (Disentangled Attention)	1
6.10	SpanBERT (Span Boundary Objective)	1
6.11	Hugging Face: `pipeline` vs `AutoModel`	1
6.12	BERTology / Probing (Attention Head · Layer)	1

📒 สัปดาห์ 7 — Decoder: GPT · รวม 11 ข้อ

#	หัวข้อย่อย	ข้อ
7.1	Causal Language Modeling (CLM) · Signal Density	2
7.2	วิวัฒนาการ GPT-1 → GPT-2 → GPT-3 · In-Context Learning	1
7.3	Modern Causal Stack (ภาพรวม) · Open Weights vs Closed Source	1
7.4	RoPE (Rotary Position Embedding)	1
7.5	RMSNorm	1
7.6	SwiGLU	1
7.7	MQA / GQA และผลต่อขนาด KV Cache	2
7.8	Scaling Laws (Kaplan vs Chinchilla)	1
7.9	กลไกเบื้องหลัง In-Context Learning · Omnimodal · Context Extension	1

ลักษณะของคำถามที่ต้องระวัง

รูปแบบ	จำนวน	คำแนะนำ
"ข้อใดต่อไปนี้ถูกต้อง"	~11 ข้อ	ต้องอ่านและประเมิน **ครบทั้ง 5 ตัวเลือก**

รูปแบบ	จำนวน	คำแนะนำ
"ข้อใด *ไม่ถูกต้อง*"	~11 ข้อ	หาข้อที่ **ผิด** ไม่ใช่ข้อที่ถูก — พลาดกันบ่อยที่สุด
ข้อคำนวณ	~18 ข้อ	ใช้เครื่องคิดเลขได้ · **ตอนที่ 2 เกือบทั้งหมดเป็นข้อคำนวณ**

💡 คำแนะนำในการเตรียมตัว

#	คำแนะนำ
1	**ข้อสอบเน้น "เข้าใจและคำนวณ" มากกว่า "ท่องจำ"** — เนื่องจากอนุญาตให้นำชีต A4 เข้าห้องสอบ ข้อสอบจึงลดคำถามแบบท่องนิยาม/ชื่อโมเดลลง และเพิ่มคำถามที่ต้องประยุกต์สูตร คำนวณหลายขั้น และวินิจฉัยว่าข้อความใดผิดแทน → ชีต A4 ควรใช้เก็บ **สูตรและตัวอย่างการแทนค่า** มากกว่าการลอกนิยาม
2	**ฝึกคำนวณจริงจากตัวอย่างในสไลด์และ lab** — ข้อคำนวณราว 18 ข้อกระจายอยู่ทุกสัปดาห์ ควรฝึกแทนค่าเองให้คล่อง ไม่ใช่แค่อ่านผ่าน
3	**หัวข้อที่ออกหนักเป็นพิเศษ** — ⭐ KV Cache (สัปดาห์ 5) 5 ข้อ · LSTM Gates (สัปดาห์ 2) 3 ข้อ · Subword Tokenization (สัปดาห์ 1) 3 ข้อ
4	**แยกให้ออกระหว่างโมเดล/กลไกที่คล้ายกัน** — เนื้อหาหลายสัปดาห์มีคู่ที่ชื่อใกล้เคียงกันและมักถูกนำมาเป็นตัวเลือกลวง ควรทบทวนให้แยกออกชัดเจน
5	**อ่านจากสไลด์เป็นหลัก** — ข้อสอบทุกข้ออ้างอิงจากสไลด์ที่สอนในห้อง (รวมเนื้อหาใน demo / lab notebook บางส่วน) ไม่มีข้อใดที่ต้องอ่านนอกเหนือจากนั้น

ขอให้นักศึกษาทุกคนโชคดีในการสอบครับ