

คู่มือการประยุกต์ใช้ **AI** ตรวจสอบข้อสอบปรนัยและวิเคราะห์ผลสัมฤทธิ์ทางการเรียนด้วย **NotebookLM**

ผู้เขียน: อาจารย์ ดร.ไชย มีหนองหว้า | สังกัด: มหาวิทยาลัยราชภัฏวไลยอลงกรณ์ ในพระบรมราชูปถัมภ์ | เผยแพร่: www.rajabhat.com

วัตถุประสงค์และประโยชน์หลักของคู่มือ

คู่มือฉบับนี้จัดทำขึ้นเพื่อให้คณาจารย์และบุคลากรทางการศึกษาสามารถนำเทคโนโลยีปัญญาประดิษฐ์ (AI) บน Google NotebookLM มาใช้ตรวจสอบข้อสอบปรนัยจากภาพถ่ายกระดาษคำตอบ เปรียบเทียบกับเฉลยมาตรฐาน สรุปผลคะแนนรายคนและทั้งชั้นเรียน ส่งออกไฟล์ Excel อัตโนมัติ พร้อมทั้งวิเคราะห์ข้อสอบ (Item Analysis) และการตัดเกรดอิงกลุ่มด้วยคะแนนมาตรฐาน T-Score อย่างถูกต้องตามหลักวิชาการ

1. บทนำและข้อได้เปรียบของการตรวจสอบข้อสอบด้วย AI

ในการจัดการเรียนการสอนระดับอุดมศึกษาที่มีนักศึกษาจำนวนมาก การตรวจสอบข้อสอบปรนัยด้วยแผ่นทาบเจาะรูใช้เวลา นานและเสี่ยงต่อความผิดพลาด ขณะที่เครื่องตรวจ OMR Scanner มีราคาสูง การประยุกต์ใช้ NotebookLM ซึ่งมีเทคโนโลยี Multimodal AI จึงช่วยเพิ่มประสิทธิภาพได้ดังนี้:

ประเด็นการประเมิน	การตรวจแบบดั้งเดิม (แผ่นทาบ)	การใช้ AI ตรวจสอบ (NotebookLM)
ความเร็วในการตรวจ	เฉลี่ย 2-3 นาที / แผ่น (40 คน ใช้เวลา 2 ชม.)	เฉลี่ย 10-15 วินาที / แผ่น (40 คน ใช้เวลา 5-10 นาที)
ความแม่นยำและการตรวจสอบ	มักตาลาย เว้นว่างหลุดสายตา หรือ รวมคะแนนผิด	ตรวจเทียบพิกัดแบบดิจิทัล มีบันทึกเปรียบเทียบรายข้อ
การออกรายงานสรุป	ต้องนำคะแนนไปพิมพ์กรอกลง Excel เองทีละคน	ระบบสร้างไฟล์ Excel พร้อมสูตรคำนวณและสถิติเสร็จสรรพ
การประเมินผลการเรียน	ตัดเกรดอิงเกณฑ์คงที่ หากข้อสอบ ยากนักศึกษาคงตกยกขึ้น	คำนวณคะแนนมาตรฐาน T-Score เพื่อตัดเกรดอิงกลุ่มได้ทันที

2. การเตรียมความพร้อมก่อนการตรวจสอบข้อสอบ (Preparation)

- ไฟล์เฉลยมาตรฐาน (Answer Key): พิมพ์ลงไฟล์ Word หรือ Docs ระบุข้อ 1 ถึงข้อสุดท้ายอย่างชัดเจน เช่น 'ข้อ 1. จ, ข้อ 2. ข, ข้อ 3. ค, ...'
- ภาพถ่ายกระดาษคำตอบ (Answer Sheets): ถ่ายภาพกระดาษคำตอบหน้าตรง แสงสว่างสม่ำเสมอ ไม่มีเงาบัง และเห็นขอบกระดาษครบทั้ง 4 ด้าน
- การตั้งชื่อไฟล์ภาพ: แนะนำให้ตั้งชื่อไฟล์ภาพด้วย 'รหัสประจำตัวนักศึกษา' (เช่น 69123020021.jpg) เพื่อความสะดวกในการจับคู่คะแนน

3. ขั้นตอนการปฏิบัติงาน 4 ขั้นตอน (Step-by-Step Workflow)

ขั้นตอนที่ 1: สร้างโครงการ (Notebook) และอัปโหลดแหล่งข้อมูล — เข้า notebooklm.google.com คลิก 'New Notebook' และอัปโหลดไฟล์เฉลยมาตรฐานพร้อมภาพถ่ายคำตอบทั้งหมดเข้าสู่กล่อง Add Sources

ขั้นตอนที่ 2: กำหนดเกณฑ์เฉลยมาตรฐานในระบบ — ส่งคำสั่งเพื่อบันทึกคีย์เฉลยที่ต้องการ:

🔗 Prompt 1 การกำหนดเกณฑ์เฉลี่ยมาตรฐาน

"ให้ตรวจสอบแหล่งข้อมูลเฉลี่ย [ระบุชื่อไฟล์เฉลี่ย เช่น สรุปเฉลยคำตอบ VGE111 ฉบับปรับปรุงใหม่ v3] และสรุปคำตอบตั้งแต่ข้อ 1 ถึงข้อ 60 ออกมาเป็นลำดับ เพื่อใช้เป็นเกณฑ์มาตรฐานในการตรวจกระดาษคำตอบของนักศึกษาทั้งหมด"

ขั้นตอนที่ 3: การตรวจข้อสอบเปรียบเทียบรายบุคคลและส่งออก Excel — ส่งคำสั่งตรวจทีละคนเพื่อตรวจสอบความถูกต้อง:

🔗 Prompt 2 : การตรวจข้อสอบรายบุคคลแบบเปรียบเทียบคำตอบ

"ทำการตรวจข้อสอบจากรูปภาพกระดาษคำตอบ [ระบุชื่อไฟล์ เช่น 69123020021.jpg]

โดยเปรียบเทียบคำตอบที่นักศึกษาระบายกับเฉลยมาตรฐานทีละข้อตั้งแต่ข้อ 1 ถึง 60:

1. แสดงสรุปคะแนนรวมที่ได้ และคิดเป็นร้อยละ
2. แสดงตารางเปรียบเทียบคำตอบแบ่งตามช่วงข้อ (1-10, 11-20, ...) ระบุข้อที่ตอบถูกและข้อที่ตอบผิดอย่างละเอียด
3. สร้างไฟล์สเปรดชีต Excel (.xlsx) สรุปผลการตรวจรายข้อของนักศึกษาคนนี้ พร้อมสูตรคำนวณคะแนนรวมอัตโนมัติ"

ขั้นตอนที่ 4: การสรุปคะแนนทั้งชั้นเรียนและการเรียงลำดับรหัส — เมื่อตรวจครบแล้ว ให้ประมวลผลทั้งห้อง:

🔗 Prompt 3 : การรวมคะแนนทั้งห้องและคำนวณสถิติภาพรวม

"จัดทำตารางสรุปคะแนนสอบของนักศึกษาทั้งหมดที่มีการตรวจไปแล้ว โดยเรียงลำดับตามรหัสประจำตัวนักศึกษาจากน้อยไปมาก:

1. แสดงคอลัมน์: ลำดับ, รหัสประจำตัว, ชื่อ-นามสกุล, คะแนนที่ได้ (เต็ม 60), ร้อยละ, เกรดอิงเกณฑ์, สถานะผ่าน/ไม่ผ่าน (เกณฑ์ 50%)
2. คำนวณสถิติภาพรวม: คะแนนเฉลี่ย (Mean), มัธยฐาน (Median), ส่วนเบี่ยงเบนมาตรฐาน (SD), คะแนนสูงสุด (Max), คะแนนต่ำสุด (Min), อัตราการสอบผ่าน
3. ส่งออกผลลัพธ์เป็นไฟล์ Excel (.xlsx) บันทึกลงใน Studio เพื่อนำไปใช้งานต่อไป"

4. การวิเคราะห์คุณภาพข้อสอบ (Item Analysis)

การวิเคราะห์ข้อสอบรายข้อช่วยให้อาจารย์ทราบค่าความยากง่าย (p) อำนาจจำแนก (r) และประสิทธิภาพของตัวลองง เพื่อใช้ทำรายงาน มคอ.5:

🔗 Prompt 4 : การวิเคราะห์ข้อสอบรายข้อและข้อบกพร่องของคำถาม

"จากข้อมูลการตรวจกระดาษคำตอบของนักศึกษาทั้งหมดในห้องนี้:

1. จัดอันดับ 5 ข้อสอบที่นักศึกษาตอบผิดมากที่สุด พร้อมระบุว่าตัวเลือกใดที่นักศึกษาส่วนใหญ่เลือกผิด
2. วิเคราะห์สาเหตุเบื้องต้นว่าเกิดจากตัวลองงมีประสิทธิภาพ คำถามกำกวม หรือนักศึกษาขาดความเข้าใจในเนื้อหาเรื่องดังกล่าว
3. สรุปเป็นตารางจำแนกสัดส่วนความยากง่ายของข้อสอบในภาพรวม"

5. การตัดเกรดอิงกลุ่มด้วยคะแนนมาตรฐาน T-Score (Norm-Referenced Grading)

เมื่อข้อสอบมีความยากสูง หากใช้การตัดเกรดแบบอิงเกณฑ์คงที่ (เช่น ผ่านที่ 50% หรือ 30 คะแนน) อาจมีนักศึกษาสอบตกเป็นจำนวนมากเกินจริง การตัดเกรดอิงกลุ่มด้วยคะแนนมาตรฐาน T-Score จึงช่วยสะท้อนความสามารถที่แท้จริงตามการกระจายตัวแบบโค้งปกติ (Normal Distribution)

☐ สูตรการคำนวณ Z-Score และ T-Score

1. คะแนนมาตรฐาน Z (Z-Score): $Z = (X - \bar{X}) / SD$ [X = คะแนนดิบ, $\bar{X}$ = คะแนนเฉลี่ย, SD = ส่วนเบี่ยงเบนมาตรฐาน]
2. คะแนนมาตรฐาน T (T-Score): $T = 50 + 10 \times Z$ [ค่าเฉลี่ย = 50, $SD = 10$ เสมอ ช่วยขจัดปัญหาคะแนนติดลบและทศนิยม]

กรณีศึกษาจริง: วิชา VGE111 ทักษะการรู้สารสนเทศ (N = 38 คน)

ผลการตรวจข้อสอบด้วย AI พบว่า: คะแนนเต็ม 60, คะแนนเฉลี่ย ($\bar{X}$) = 29.03 คะแนน, $SD = 6.29$ คะแนน, สูงสุด 47, ต่ำสุด 16

- อิงเกณฑ์คงที่ (ผ่านที่ 50% หรือ ≥ 30 คะแนน): สอบผ่าน 18 คน (47.4%) และสอบตกถึง 20 คน (52.6%)
- อิงกลุ่ม T-Score (ผ่านที่ $T \geq 35$ หรือ ≥ 20 คะแนน): สอบผ่าน 35 คน (92.1%) และตกเพียง 3 คน (7.9%) ซึ่งสะท้อนการประเมินที่เป็นธรรม

เกรด	ช่วงคะแนน T-Score	เกณฑ์เทียบส่วนเบี่ยงเบน (SD)	ช่วงคะแนนดิบโดยประมาณ (เต็ม 60)	ความหมาย
A	$T \geq 65$	สูงกว่าค่าเฉลี่ย $\geq +1.5 SD$	39 คะแนนขึ้นไป	ยอดเยี่ยม (Excellent)
B+	$60 \leq T < 65$	+1.0 SD ถึง +1.5 SD	35 – 38 คะแนน	ดีมาก (Very Good)
B	$55 \leq T < 60$	+0.5 SD ถึง +1.0 SD	32 – 34 คะแนน	ดี (Good)
C+	$50 \leq T < 55$	ค่าเฉลี่ย ถึง +0.5 SD	29 – 31 คะแนน	ดีพอใช้ (Fairly Good)
C	$45 \leq T < 50$	-0.5 SD ถึง ค่าเฉลี่ย	26 – 28 คะแนน	พอใช้ (Fair)
D+	$40 \leq T < 45$	-1.0 SD ถึง -0.5 SD	23 – 25 คะแนน	อ่อน (Poor)
D	$35 \leq T < 40$	-1.5 SD ถึง -1.0 SD	20 – 22 คะแนน	อ่อนมาก (Very Poor)
F	$T < 35$	ต่ำกว่าค่าเฉลี่ย $< -1.5 SD$	ต่ำกว่า 20 คะแนน	ตก (Fail)

🖨️ สูตร Excel สำเร็จรูปสำหรับตัดเกรด T-Score (คะแนนดิบอยู่ที่เซลล์ E6, ค่าเฉลี่ยที่ \$E\$46, SD ที่ \$E\$47)

1. สูตรคำนวณ T-Score (คอลัมน์ F): $=50 + 10 * ((E6 - \$E\$46) / \$E\$47)$
2. สูตรตัดเกรด 8 ระดับอัตโนมัติ (คอลัมน์ G):
 $=IFS(F6 \geq 65, "A", F6 \geq 60, "B+", F6 \geq 55, "B", F6 \geq 50, "C+", F6 \geq 45, "C", F6 \geq 40, "D+", F6 \geq 35, "D", TRUE, "F")$

📌 **Prompt 5 : สั่งให้ NotebookLM ตัดเกรดอิงกลุ่ม (T-Score) และเพิ่มลงใน Excel**

"นำผลการตรวจคะแนนสอบของนักศึกษาทุกคนมาคำนวณคะแนนมาตรฐาน T-Score ตามสูตร $T = 50 + 10 \cdot (X - \text{Mean}) / \text{SD}$:

1. สร้างคอลัมน์คะแนน T-Score และเกรดอิงกลุ่ม 8 ระดับ (A, B+, B, C+, C, D+, D, F)
2. ทำตารางแจกแจงความถี่จำนวนนักศึกษาในแต่ละเกรด
3. อัปเดตข้อมูลทั้งหมดลงในไฟล์ Excel สรุปผลสอบ พร้อมส่งออกให้ดาวน์โหลด"

6. แนวปฏิบัติที่ดีและข้อควรระวัง (Best Practices)

1. การตรวจสอบข้อบกพร่อง: หากนักศึกษาระบายคำตอบไม่ชัดเจนหรือเว้นว่าง ระบบจะลงเป็น (เว้น) หรือให้ 0 คะแนน อาจารย์สามารถคลิกดูภาพต้นฉบับในสมุดงานเพื่อตรวจยืนยันความถูกต้องซ้ำได้ทันที
2. ความโปร่งใสของผลการประเมิน: อาจารย์สามารถแนบตารางเปรียบเทียบรายข้อให้นักศึกษาตรวจสอบความถูกต้องของตนเองได้ ซึ่งช่วยลดข้อโต้แย้งและเพิ่มความโปร่งใสในการวัดผล
3. การสำรองข้อมูล: เมื่อสิ้นสุดภาคการศึกษา ควรดาวน์โหลดไฟล์ผลการตรวจรายคนและไฟล์สรุปภาพรวม (.xlsx) จาก แอป Studio จัดเก็บไว้ใน Google Drive ของรายวิชาเพื่อเป็นหลักฐานทางวิชาการต่อไป