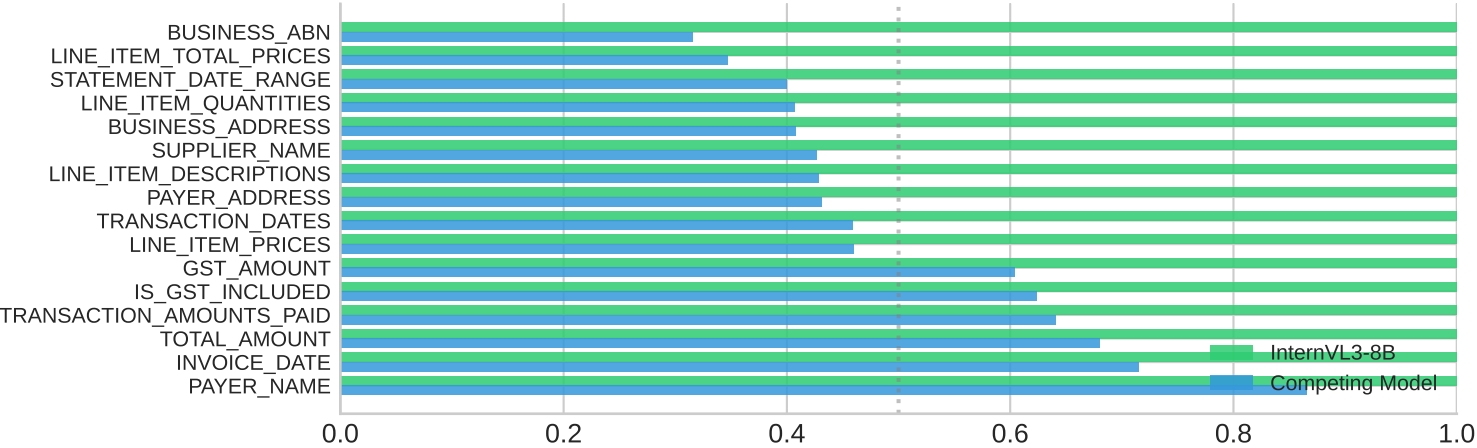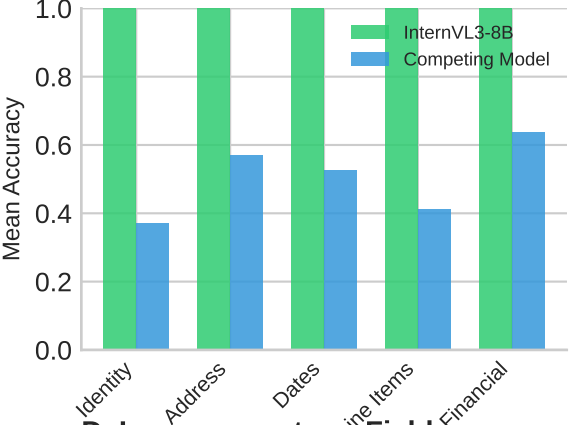# Model Accuracy Comparison Dashboard

InternVL3-8B vs Competing Model | 16 comparable fields | Mean improvement: +48.7%
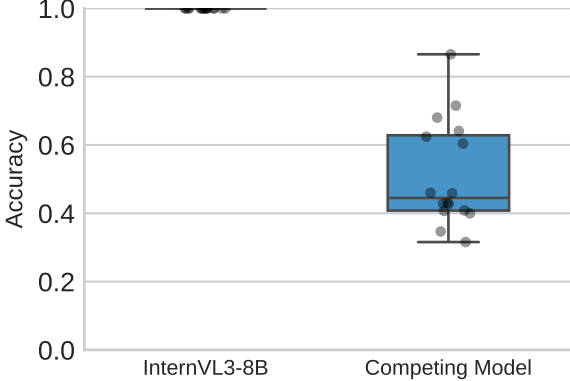
A. Field-Level Accuracy (Comparable Fields)
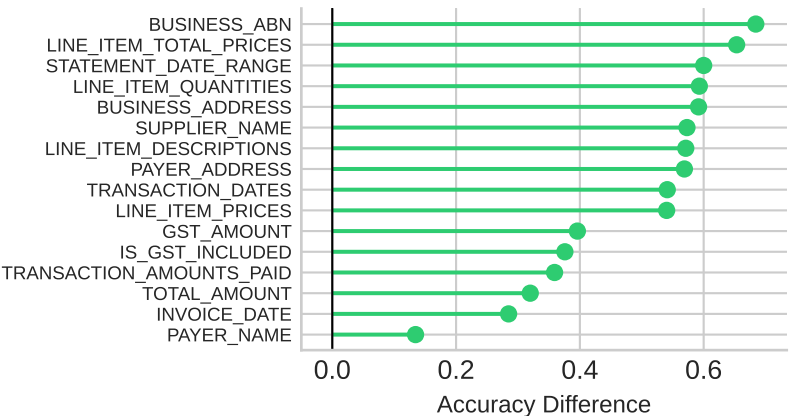

Legend: InternVL3-8B (green), Competing Model (blue)

Fields (top to bottom): BUSINESS_ABN, LINE_ITEM_TOTAL_PRICES, STATEMENT_DATE_RANGE, LINE_ITEM_QUANTITIES, BUSINESS_ADDRESS, SUPPLIER_NAME, LINE_ITEM_DESCRIPTIONS, PAYER_ADDRESS, TRANSACTION_DATES, LINE_ITEM_PRICES, GST_AMOUNT, IS_GST_INCLUDED, TRANSACTION_AMOUNTS_PAID, TOTAL_AMOUNT, INVOICE_DATE, PAYER_NAME

X-axis: Accuracy (0.0 to 1.0)

B. By Category

Y-axis: Mean Accuracy (0.0 to 1.0)
Categories: Identity, Address, Dates, Line Items, Financial
Legend: InternVL3-8B (green), Competing Model (blue)

C. Distribution

Y-axis: Accuracy (0.0 to 1.0)
Box plots for InternVL3-8B and Competing Model

D. Improvement per Field

Fields (top to bottom): BUSINESS_ABN, LINE_ITEM_TOTAL_PRICES, STATEMENT_DATE_RANGE, LINE_ITEM_QUANTITIES, BUSINESS_ADDRESS, SUPPLIER_NAME, LINE_ITEM_DESCRIPTIONS, PAYER_ADDRESS, TRANSACTION_DATES, LINE_ITEM_PRICES, GST_AMOUNT, IS_GST_INCLUDED, TRANSACTION_AMOUNTS_PAID, TOTAL_AMOUNT, INVOICE_DATE, PAYER_NAME

X-axis: Accuracy Difference (0.0 to 0.6)

E. Summary Statistics

Metric	InternVL3-8B	Competing Mo
Mean	100.0%	51.3%
Median	100.0%	44.5%
Std Dev	0.0%	15.5%
Critical Fields	100.0%	50.7%
Statistical Test	Value	Result
p-value (t-test)	0.0000	Significant
Cohen's d	4.60	large
95% CI	[+41.1%, +55.5%]	

F. ADDITIONAL CAPABILITIES (InternVL3-8B only)

The following fields are unique to our model (not available in competing model):

• DOCUMENT_TYPE: 100.0% accuracy

Mean accuracy of exclusive capabilities: 100.0%