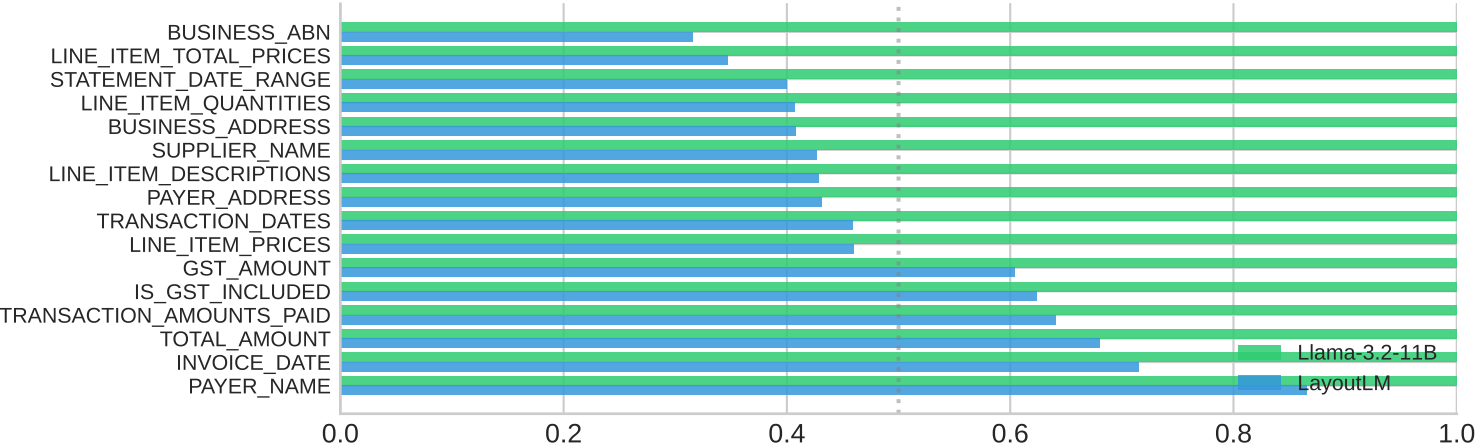


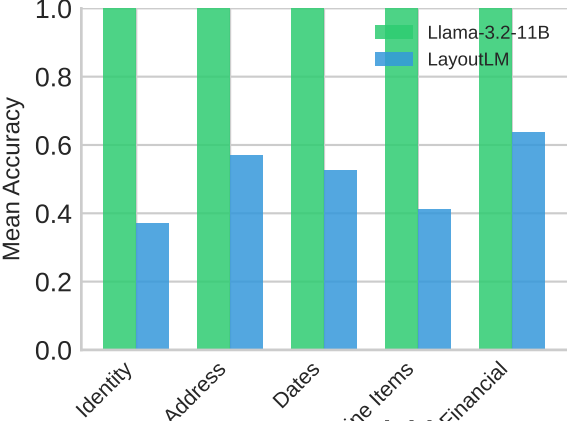
# Model Accuracy Comparison Dashboard

Llama-3.2-11B vs LayoutLM | 16 comparable fields | Mean improvement: +48.7%

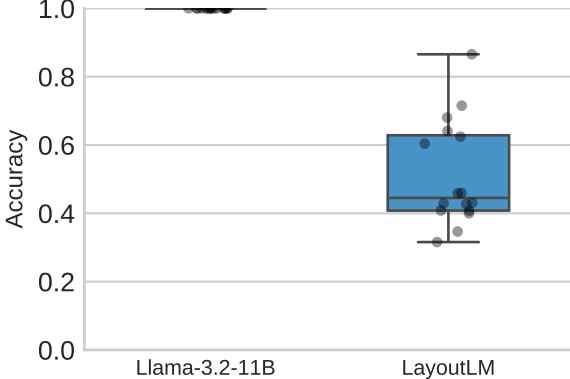
## A. Field-Level Accuracy (Comparable Fields)



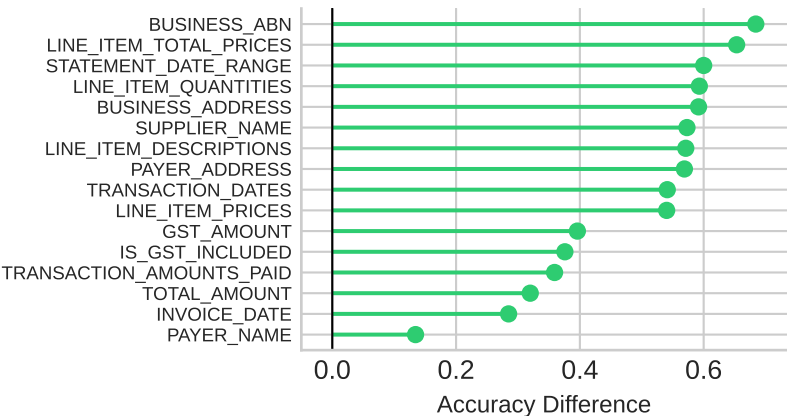
## B. By Category



## C. Distribution



## D. Improvement per Field



## E. Summary Statistics

Metric	Llama-3.2-11	LayoutLM
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	[-41.0%, +55.6%]	

## F. ADDITIONAL CAPABILITIES (Llama-3.2-11B 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%