

 Guide 1, 15 minutes, 4 steps

# A business unit's first AI model

From one task to a pilot you can measure. For the people who run a process, not for the people who train models. This PDF is the printed version of an interactive guide. To run the experiments yourself: [halilneed.agency/guides/first-model/](https://halilneed.agency/guides/first-model/)

# Which task first?

Pick a task. The task card fills itself in.



Task card

✉ Sort emails

📄 Pull fields from invoices

📄 Summarise requests

👤 Approve credit applications



Input

An incoming customer email



Output

One of six categories



Examples

50 or more labelled emails per category



Measure of success

Share in the right category, per category



Human check

When the model is unsure, a person takes it

✓ Repeats often

✓ Easy to check

✓ Cheap to get wrong

✓ A good first task



A good first task repeats often, has an output you can check and is cheap to get wrong.

## One score hides the weak cases

Look at the overall score, then split it into slices.

# 94%

Overall accuracy

Overall score

Split into slices

### Short emails

50% of the data



97%

### Long emails

25% of the data



94%

### Emails with typos

10% of the data



88%

### Emails in English

8% of the data



89%

### Rare category: refunds

7% of the data



86%

The figures are illustrative. I took my own Turkish personal-data model from 0.740 to 0.882 this way, without training a bigger model.

💡 Not a bigger model: targeted examples for the weak slices.

## Which level to start at?

Answer three questions and the starting level lights up.



### Steps 3 and 4 are in the full guide

The full guide is on Gumroad as a PDF in English and Turkish. It is free; pay what you want if you like.

- ✓ Step 3 of 4: Which level to start at?
- ✓ Step 4 of 4: Shadow first, then assist
- ✓ The full guide as a PDF, English and Turkish

Steps 3 and 4 and the full PDF are free. To get them: [halilneed.gumroad.com/l/ilik-model](https://halilneed.gumroad.com/l/ilik-model)