CATER EXTRACT

Presented by:
Arkadi Doktorovich, Hen Mandelbaum

Motivation

- Business agreements are often finalized through email thread negotiations, Important contractual details are embedded implicitly across several messages making manual extraction time consuming and error prone.

- The goal of this project is to automatically extract structured contract attributes from catering email threads and output them in a fixed JSON schema.

Project definition

- Input: Multi-turn catering email threads
- Output: Fixed JSON contract schema
- For models we used a LLM that uses a prompt (GPT 4-o-mini for zero-shot and few-shot), supervised Transformer (DistilBERT multi-head).
- For the Dataset we created 1,000 synthetic but realistic email threads with aligned JSON ground truths.
- The metrics used are Field level exact match / overlap scores, regression tolerance accuracy, multi label F1, and overall average score.

Email 1 - Food Supplier:
Subject: Catering Proposal for Upcoming Event

Dear Sarah,

I hope this message finds you well. Thank you for reaching out to us at Gourmet Delights regarding your upcoming event. We would be delighted to cater for your corporate gathering scheduled for March 15th.

Our pricing structure is typically set at $45 per person, which includes a variety of dishes tailored to your needs. We can accommodate a minimum of 50 guests and a maximum of 150 guests. All our food is prepared under strict kosher supervision by the OK, ensuring the utmost quality and compliance.

Please note that beverages are not included in our standard package, but we can arrange a bar service at an additional cost. We can also offer a mixed menu featuring options like grilled chicken, roasted vegetables, and a selection of pareve desserts.

I would be happy to provide a sample menu for your review and discuss any dietary accommodations you may require.

Looking forward to your response.

Best regards,
David Rosen
Gourmet Delights

Email 2 - Event Organizer:
Subject: Re: Catering Proposal for Upcoming Event

Hi David,

Thank you for your prompt response and for outlining the details. The $45 per person rate seems reasonable, but I was hoping to discuss the possibility of a fixed price for the entire event if we confirm a higher number of guests, say around 120. Could you provide a quote based on that?

Additionally, it would be great to know more about your vegetarian options, as we may have some guests with dietary restrictions. The event will be a formal networking dinner, so presentation is key.

Thank you, and I look forward to hearing back from you soon.

Best,
Sarah Thompson
Event Horizon

Email 3 - Food Supplier:
Subject: Re: Catering Proposal for Upcoming Event

Dear Sarah,

Thank you for your reply. I appreciate your interest in a fixed price arrangement. For 120 guests, we can offer a package deal at $5,400, which includes all the menu items and necessary staffing for setup and service. This price does include VAT.

Regarding dietary accommodations, we can certainly provide vegetarian options, like a quinoa salad and stuffed peppers, alongside our meat dishes. I assure you that we will handle all presentation details to match the formal setting of your event.

Please let me know if this arrangement works for you, and I'll send over the formal agreement.

Best wishes,
David Rosen
Gourmet Delights

Email 4 - Event Organizer:
Subject: Re: Catering Proposal for Upcoming Event

Hi David,

Thank you for the revised proposal. The fixed price of $5,400 for 120 guests works for us, and I appreciate your willingness to accommodate our dietary needs. I'm happy to proceed with the agreement.

Please send over the formal contract at your earliest convenience, along with details regarding the cancellation policy and any other relevant notes for the event.

Thank you for your assistance, and I look forward to finalizing this.

Best,
Sarah Thompson
Event Horizon

```
{
  "event_type": "corporate gala",
  "price_type": "per person",
  "final_price_value": 9750,
  "min_guests": 50,
  "max_guests": 200,
  "includes_vat": true,
  "is_kosher": true,
  "kosher_supervision": "Orthodox Union",
  "includes_bar": true,
  "menu_type": "mixed",
  "dietary_options": "vegetarian, gluten-free",
  "event_date": "March 15",
  "cancellation_policy": "14 days for full refund, 50% charge after",
  "menu_highlights": "herb-marinated chicken, roasted seasonal vegetables, desserts",
  "extra_notes": "setup and staffing included"
}
```

```
{
  "event_type": "gala dinner",
  "price_type": "fixed price",
  "final_price_value": 3900,
  "min_guests": 50,
  "max_guests": 150,
  "includes_vat": true,
  "is_kosher": true,
  "kosher_supervision": "Orthodox Union",
  "includes_bar": false,
  "menu_type": "mixed",
  "dietary_options": "vegetarian, gluten-free",
  "event_date": "2023-12-10",
  "cancellation_policy": "14 days",
  "menu_highlights": [
    "herb-crusted salmon",
    "roasted vegetable lasagna",
    "chicken piccata",
    "seasonal vegetable stir-fry"
  ],
  "extra_notes": "setup timing and equipment provision needed"
}
```

Feature Schema

Feature Name	Outcome Type	Possible Values / Description
event_type	Categorical (Text)	String describing event type or null
price_type	Categorical (Enum)	per_person, fixed, range, unknown
final_price_value	Numerical	Price value or null
min_guests	Numerical	Minimum number of guests or null
max_guests	Numerical	Maximum number of guests or null
includes_vat	Boolean	true, false, or null
is_kosher	Boolean	true, false, or null
kosher_supervision	Categorical (Text)	Supervision authority or null
includes_bar	Boolean	true, false, or null
menu_type	Categorical (Enum)	meat, dairy, mixed, pareve, unknown
dietary_options	Multi-Label Categorical	Array of dietary strings (e.g., vegan, gluten-free)
event_date	Temporal (Date)	YYYY-MM-DD or null
cancellation_policy	Categorical (Text)	Policy description or null
menu_highlights	Multi-Label Categorical	Array of highlighted menu items
extra_notes	Free Text	Additional notes or null

Feature Negotiation

Our task is not to extract mentions, but to extract the final agreed values.
Here is an example of values where the values change as the negotiation progresses.

Email #	Sender	Price Type	Price Value	Guests	Includes Bar	Dietary Options	Cancellation Policy
1	Supplier	Per Person	$75	50–150	No	Vegetarian	Not specified
2	Organizer	Per Person	Request discount	100	Asked	Vegetarian, Gluten-Free	Asked
3	Supplier	Per Person	$70	100	Optional	Vegetarian, Gluten-Free	30 days full refund
4	Organizer	Per Person	$70	100	Yes	Vegetarian, Gluten-Free	Accepted

Data Creation Pipeline

Goal: Generate realistic business email negotiations with structured ground truth.

- Synthetic email threads were generated using a large language model to simulate negotiations between a food supplier and an event organizer.
- Each thread contains **four emails**, including proposal, negotiation, clarification, and final agreement or cancellation.
- Information is embedded **implicitly** in natural language, without explicit field labels, to mimic real world communication.
- For every thread, the model outputs a **structured JSON ground truth** containing event attributes (e.g., price, guests, kosher status, cancellation policy).
- Missing information is explicitly represented as **null**, enabling realistic partial information scenarios.
- A total of **1,000 email threads** were generated, split into email text files and corresponding JSON labels.

Outcome: A controlled yet realistic dataset for structured information extraction from negotiation style emails.

Multi-Task DistilBERT Feature Extractor

Model Architecture:

- **Backbone:** DistilBERT encoder for contextual text representation.
- Multi-head output layers for different prediction types:
  - *Classification (event type, price type, menu type, cancellation policy)*
  - *Binary prediction (VAT, kosher, bar inclusion)*
  - *Regression (final price, guest counts, event date)*
  - *Multi-label classification (dietary options, menu highlights, extra notes)*
- Masked mean pooling over token embeddings to obtain document level representation.

Training Setup:

- Multi-task learning with **task specific loss functions** (cross-entropy, MSE, BCE).
- Field specific loss weighting to balance heterogeneous outputs.
- Numeric fields normalized to stabilize training.

Achievements and Novelty

Accomplishments:

- Trained a supervised extraction model that outperformed zero-shot and few-shot LLMs on structured feature extraction.
- Finding a close to perfect prompt for business email threads generation.

Novelty:

- Focus on negotiation aware information extraction rather than single email or sentence level extraction.
- Direct comparison between LLM based extraction and supervised multi-task learning under the same schema.

Methodology

Project process:

- Defined a structured schema and aligned all models to the same output format.
- Trained, evaluated, and compared multiple extraction approaches.

Data & Models:

- Prepared a dataset of 1000 synthetic email thread negotiations with JSON ground truth.
- Evaluated:
  - GPT 4-o-mini, LLM zero-shot & few-shot prompting.
  - Zero-Shot : use Only instructions and schema.
  - Few- shot : Use One full example Email thread and JSON output.
  - DistilBERT, Supervised Transformer based model.
- Metrics: field-level Precision, Recall, and F1-score.

Efforts:

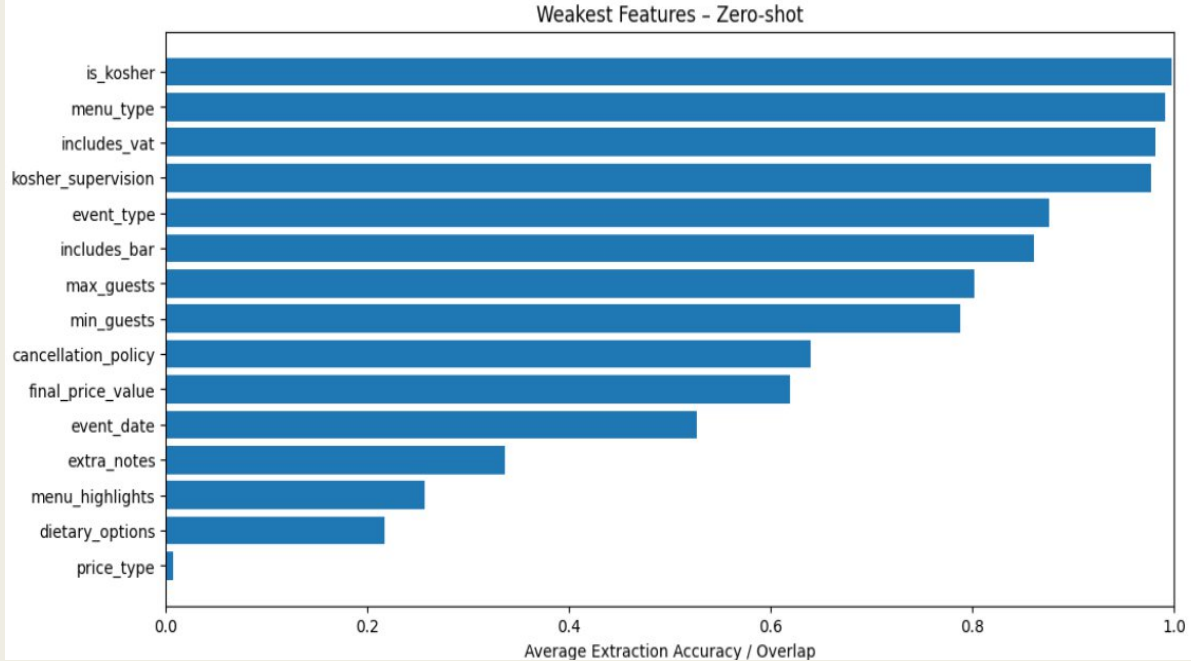
- Training from ground zero the Supervised Transformer based model.
- Prompt engineering and consistency checks for LLM outputs and data creation.

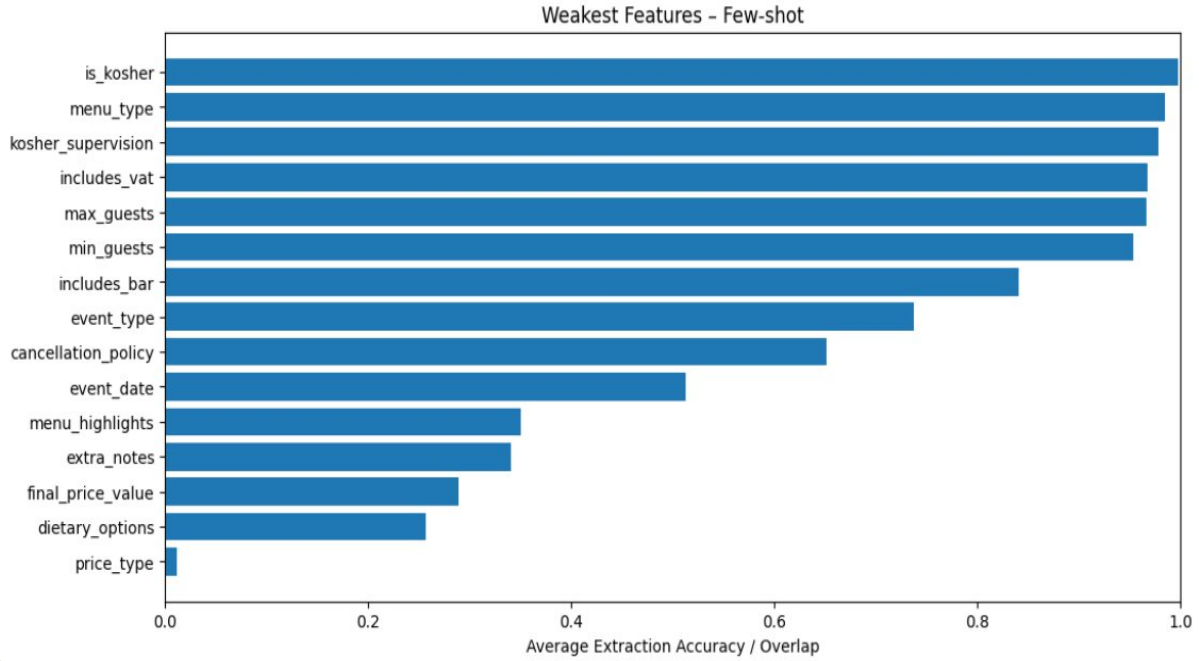
Results

```
=== Zero-shot OVERALL PERFORMANCE ===
Average exact match rate: 0.6583

=== Zero-shot FEATURE-LEVEL PERFORMANCE ===
event_type                → 0.876
price_type                → 0.007
final_price_value         → 0.619
min_guests                → 0.788
max_guests                → 0.802
includes_vat              → 0.981
is_kosher                 → 0.997
kosher_supervision        → 0.977
includes_bar              → 0.861
menu_type                 → 0.991
dietary_options           → 0.217
event_date                → 0.527
cancellation_policy       → 0.639
menu_highlights           → 0.257
extra_notes               → 0.336
```

```
Average Score (all fields): 0.7414
Per-field Scores:
  event_type              : 0.579
  price_type              : 0.944
  kosher_supervision      : 1.000
  menu_type               : 1.000
  cancellation_policy     : 0.996
  includes_vat            : 0.851
  is_kosher               : 0.998
  includes_bar            : 0.600
  final_price_value       : 0.250
  min_guests              : 0.173
  max_guests              : 0.521
  event_date              : 0.209
  dietary_options         : 1.000
  menu_highlights         : 1.000
  extra_notes             : 1.000
```

```
=== Few-shot OVERALL PERFORMANCE ===
Average exact match rate: 0.6558

=== Few-shot FEATURE-LEVEL PERFORMANCE ===
event_type                → 0.737
price_type                → 0.011
final_price_value         → 0.289
min_guests                → 0.953
max_guests                → 0.966
includes_vat              → 0.967
is_kosher                 → 0.998
kosher_supervision        → 0.978
includes_bar              → 0.841
menu_type                 → 0.985
dietary_options           → 0.257
event_date                → 0.513
cancellation_policy       → 0.651
menu_highlights           → 0.350
extra_notes               → 0.341
```





Conclusion

Our goals were to outperform the LLM models, and we successfully did it after working on the supervised transformer-based model.

Things we learnt:

- Schema design and annotation consistency are as critical as model choice.
- Prompting alone is insufficient for high precision structured extraction.

Additional ideas:

- Hybrid approaches combining LLM reasoning with supervised extraction.
- Expand dataset size and include real world negotiation data.