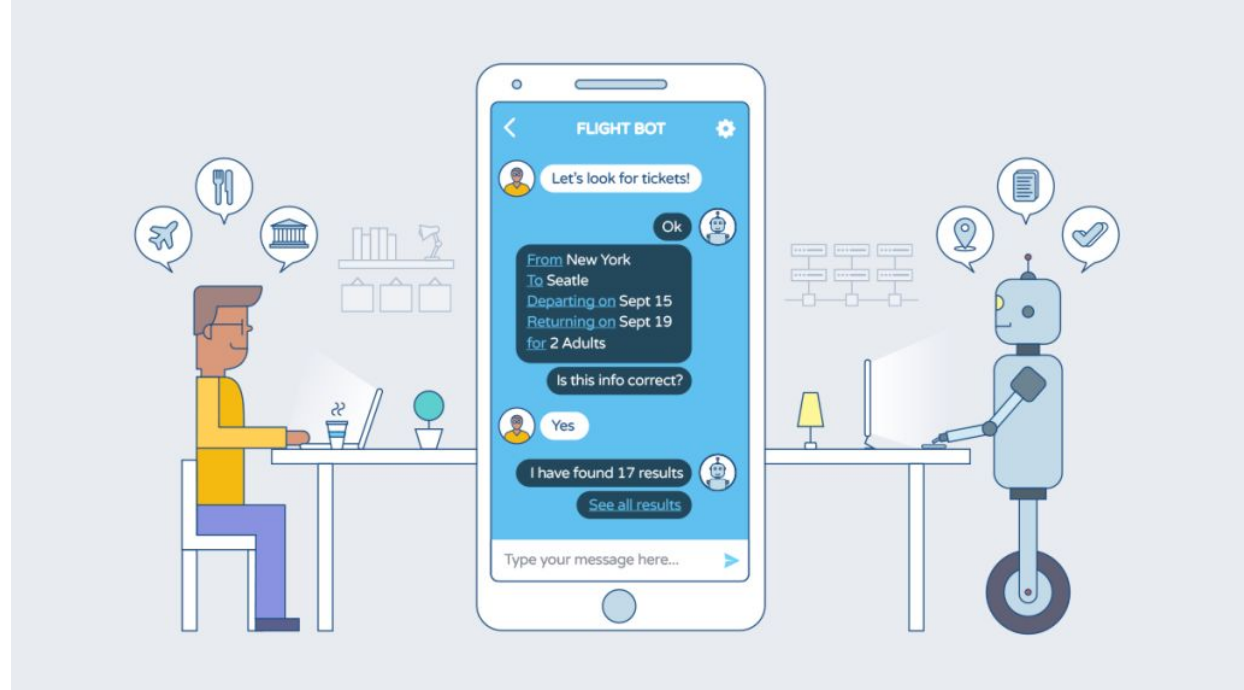


Just To Be Sure: Building A Virtual Assistant With Active Listening Skills

Piper Armstrong
Seong-Eun Cho
Yeganeh Nasiri
Aadesh Neupane
Jacob Stern



Problems



- Chatbots either overconfident or ignorant
- Voice assistants fall short in two areas: accuracy and specificity



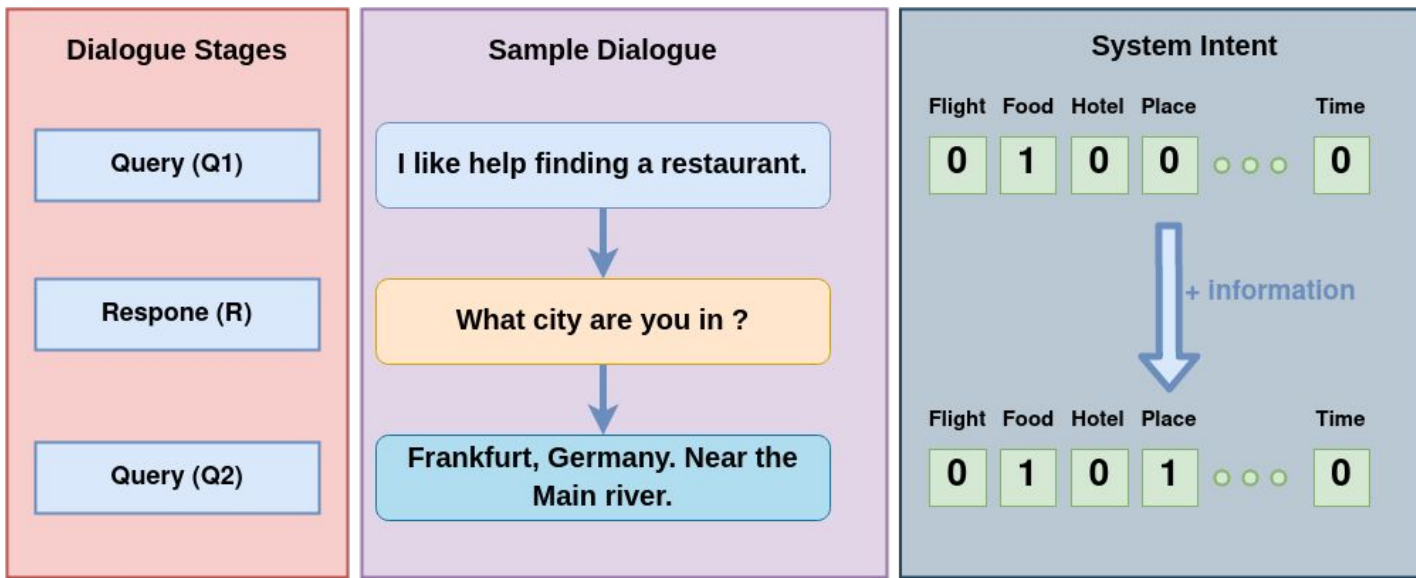
Related Works

- Existing apps and products
 - Personal assistants [Liao et al., 2018]
 - Interviewer chatbots [Li et al., 2017]
 - Mental health supporter [Inkster et al., 2018]
 - AI companions [Ta et al., 2020]
- Uncertainty
- Natural language generation and language models
 - Google Dialogflow [Sabharwal and Agrawal, 2020]
 - T5 [Bird et al., 2020]



Proposed Solution

- Identifying the source of uncertainty is vital
- Two-turn query-response-query (QRQ) task
- Ambiguous response to specify and clarifying response





Dataset

Taskmaster 2

- ❖ 17,289 dialogues
- ❖ 341,801 utterances
- ❖ 7 domains
- ❖ Fully annotated using 90 possible annotations.

"conversation_id": "dlg-0018b735-34fe-4c43-b744-e955b6836c43", "instruction_id": "movie-26"		
Annotation	Speaker	Text
movie_search.location.theater	Assistant	How can I help you?
	User	What movies are playing?
	Assistant	I can help you with your movie search.
	Assistant	Where are you located?
	User	Davis, California.
movie_search.genre	Assistant	Do you have a type of movie you wanted to go see?
	User	Action
movie_search.name.movie	Assistant	Just one moment
	Assistant	Here is what I found
	Assistant	Baby driver.
	Assistant	The mummy
	Assistant	Guardians of the Galaxy: Vol 2.
	User	The mummy
movie_search.time.start	Assistant	Okay. The Mummy is playing at 4:30 pm this afternoon at Regal Davis Stadium 5.
movie_search.time.start	User	Ok
	Assistant	Enjoy your movie!
	User	Thanks!



Task

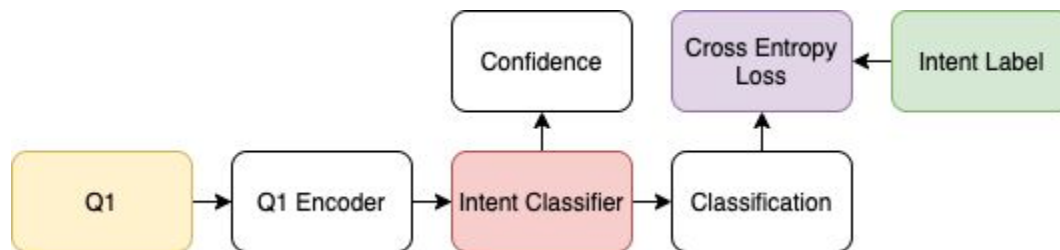
- ❖ **Query → Intent**
 - Intent Classifier
- ❖ **Query + Query → Intent**
 - Intent Classifier
- ❖ **Query → Response**
 - Encoder-Decoder
- ❖ **Query + Response → Query**
 - QRQ Model
- ❖ **Query**
 - IA Model

"conversation_id": "dlg-0018b735-34fe-4c43-b744-e955b6836c43", "instruction_id": "movie-26"		
Annotation	Speaker	Text
movie_search.location.theater	Assistant	How can I help you?
	User	What movies are playing?
	Assistant	I can help you with your movie search.
	Assistant	Where are you located?
	User	Davis, California.
movie_search.genre	Assistant	Do you have a type of movie you wanted to go see?
	User	Action
movie_search.name.movie	Assistant	Just one moment
	Assistant	Here is what I found
	Assistant	Baby driver.
	Assistant	The mummy
	Assistant	Guardians of the Galaxy: Vol 2.
	User	The mummy
movie_search.time.start	Assistant	Okay. The Mummy is playing at 4:30 pm this afternoon at Regal Davis Stadium 5.
movie_search.time.start	User	Ok
	Assistant	Enjoy your movie!
	User	Thanks!



Methods

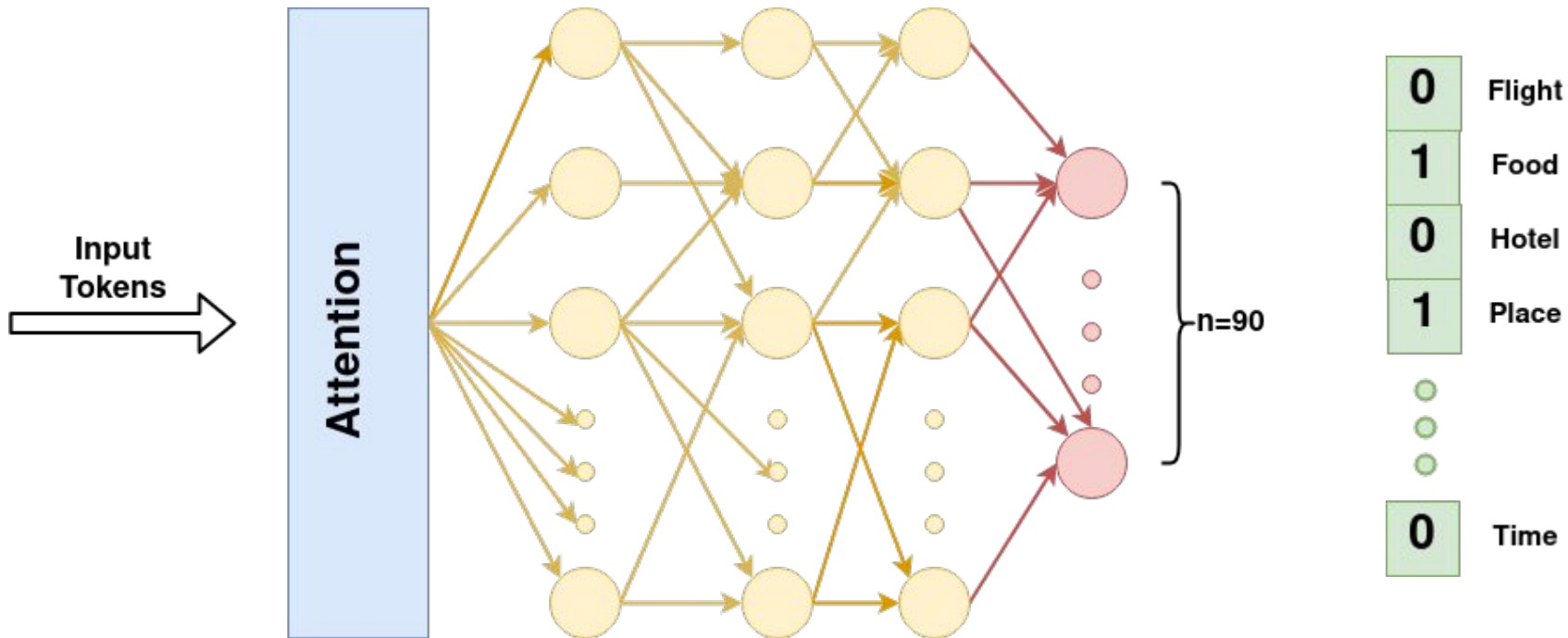
Intent Classifier





Experiments

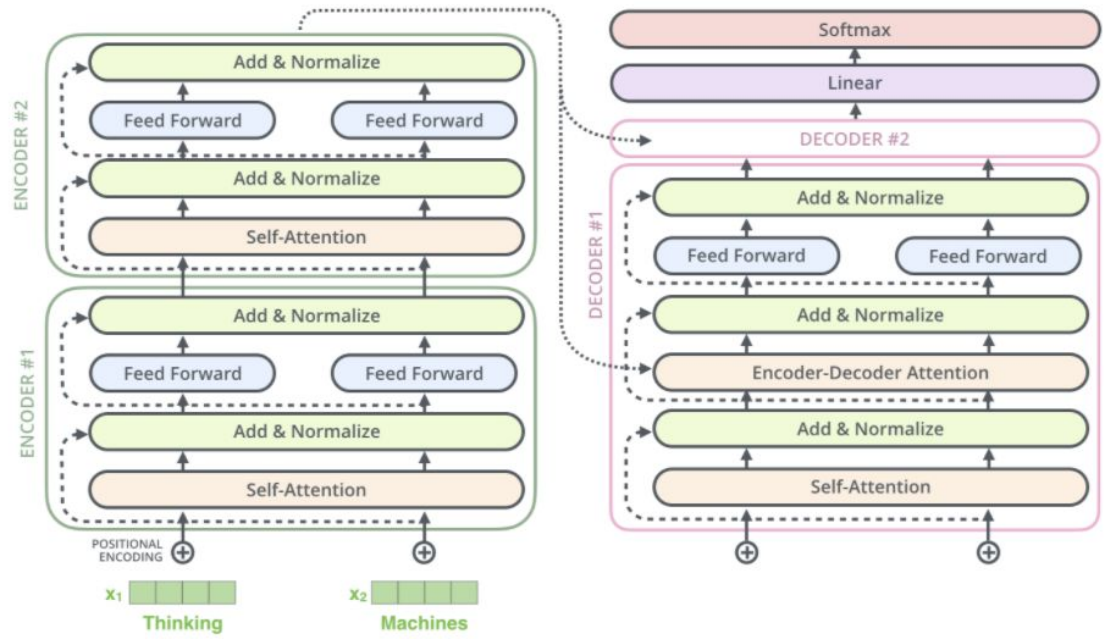
Intent Classification





Experiments

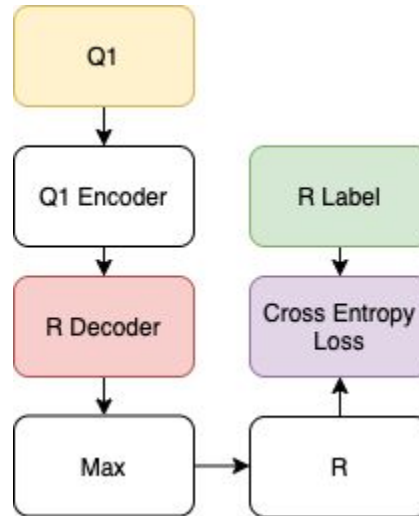
T5 Fine-tuning





Methods

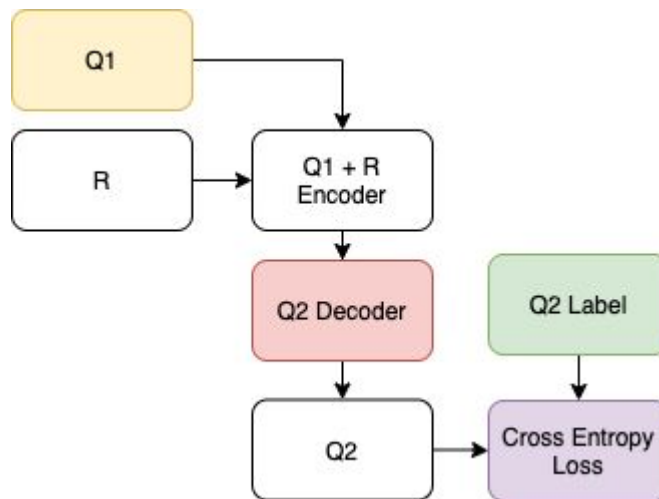
Encoder-Decoder/QR Model





Methods

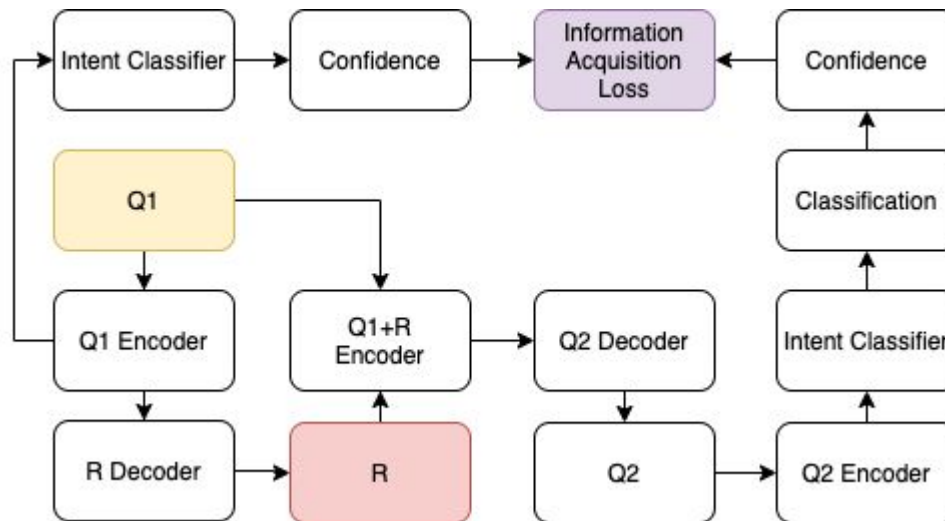
Encoder-Decoder/QRQ Model





Methods

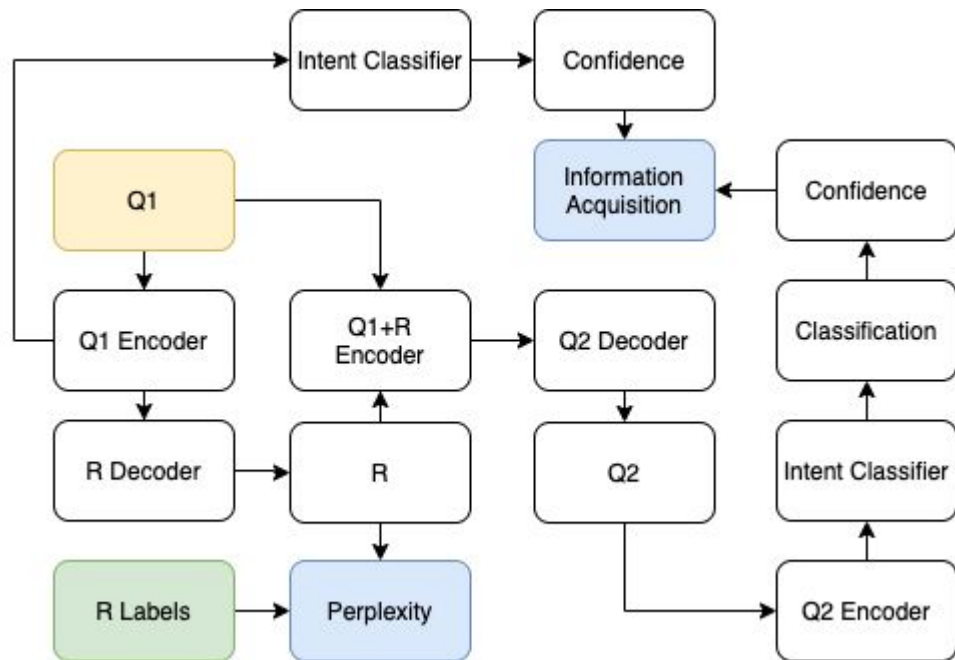
Information Acquisition (IA) Task





Methods

Evaluation of the models





Results

	F1 Score	Accuracy
QQ Intent Classifier	0.94	0.95
QQ Intent Classifier (Baseline)	0.21	0.24



Results

Evaluation metrics for our different models

	Perplexity	Information Acquisition
QR before fine-tuning (Baseline)	5.28e10	.1127
QR Transformer (Supervised)	2.32390	.1128
info acq Transformer (Unsupervised)	inf	.0977
info acq Transformer with QR fine-tuning (Supervised/Unsupervised)	inf	.095

Sample Outputs

Baseline

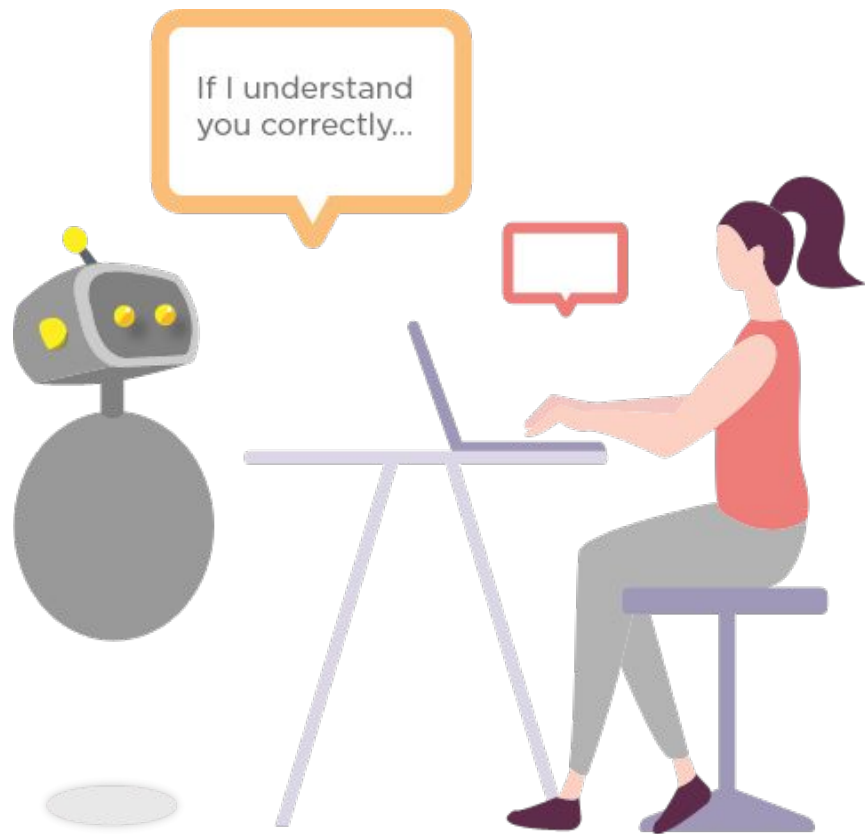
SN	Label	Text
1	Input text	I would like you to help me finding a hotel in San Francisco , Downtown area , California .
1	Output text	Ich möchte Sie bitten , mir bei der Suche nach einem Hotel in San Francisco , Downtown area , California zu helfen.
1	Target Text	What is the price range ?

Fine-tuned

SN	Label	Text
1	Input text	I'd like to see the Edge of Tomorrow trailer
1	Output text	Ok, I have your personal details on file. Your order will be ready for pickup in 20 minutes.
1	Target Text	Anything else I can help you with?

Conclusion, Future Work

- Information discernment
- Open-domain intent classes
- Open-domain action space





More Sample Generations (QRQ)

Q1: How about the 7:00 p . m ?

R: The return flights are at 12 p . m , 4:05 p.m, 11:35 p.m.

Q2 output: How about the 4:05 p.m?

Q2 target: Let's do the 4:05 p.m. return on the flight. That would be nice.

Q1: Does that include the return flight?

R: How does 5:30 a.m. work ?

Q2 output: Well, 5:30 a.m. works.

Q2 target: Well, the evening works better.

Q1: And I would like to fly during the night time and a red-eye.

R: Alright. What else ?

Q2 output: I would like to fly in the morning and I would like to fly in the morning.

Q2 target: I would like it to be under \$500.

Any query's?

