Efficient Online ML API Selection for Multi-Label Classification Tasks

ICML 2022
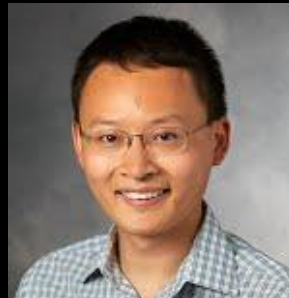
Lingjiao Chen

Joint work with

Matei Zaharia   **James Zou**

Machine Learning as a Service (MLaaS)

Machine Learning as a Service (MLaaS)

- **Goal:**

Machine Learning as a Service (MLaaS)

- Goal:

Mitigate low level overheads

Machine Learning as a Service (MLaaS)

- Goal:

Mitigate low level overheads

- e.g., model training

Machine Learning as a Service (MLaaS)
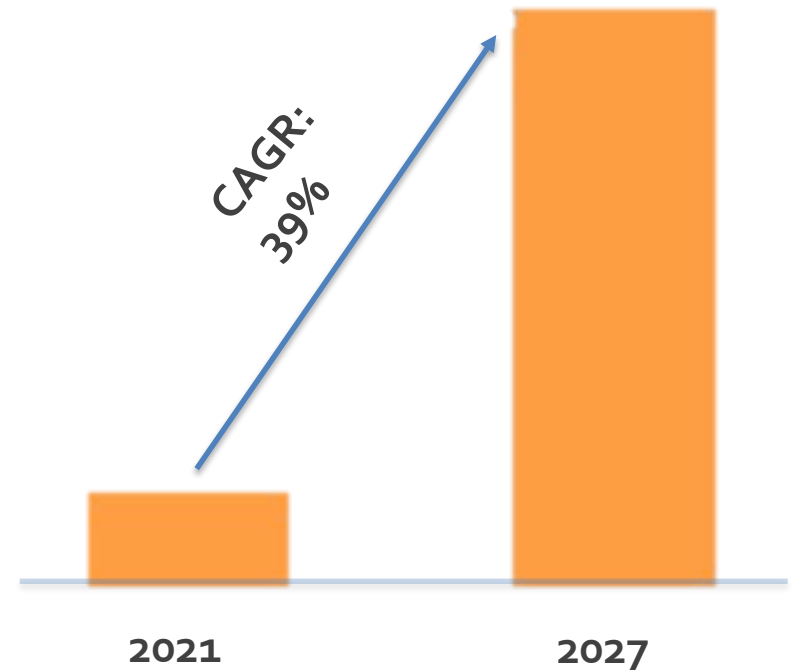
- Goal:

Mitigate low level overheads

- - e.g., model training

- - data labelling, etc

Machine Learning as a Service (MLaaS)

- Goal:

Mitigate low level overheads

- e.g., model training

- data labelling, etc

CAGR: 39%

2021   2027

Source: Mordor Intelligence

Machine Learning as a Service (MLaaS)
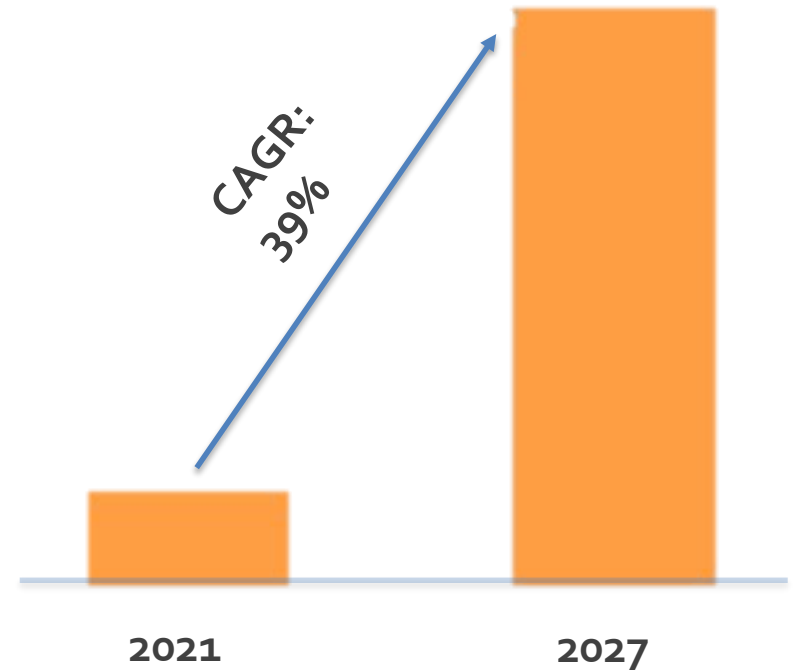
- Goal:

Mitigate low level overheads

- e.g., model training

- data labelling, etc

- VALUE:

Previous: USD 2.26 billion in 2021

Expected: USD 16.7 billion by 2027

CAGR: 39%

2021   2027

Source: Mordor Intelligence

- A common focus: multi-label classification tasks

  - Image tagging, text recognition, named entity recognition, …

…   …

- A common focus: multi-label classification tasks
    - Image tagging, text recognition, named entity recognition, ...
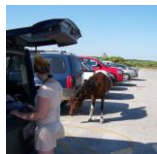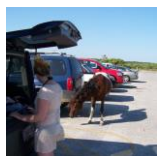- Many multi-label APIs exist

...   ...

- A common focus: multi-label classification tasks
  - Image tagging, text recognition, named entity recognition, ...
- Many multi-label APIs exist

- A common focus: multi-label classification tasks
  - Image tagging, text recognition, named entity recognition, ...
- Many multi-label APIs exist
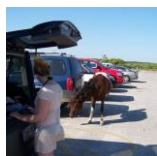- Heterogeneity in those APIs' performance and cost

...

...

- A common focus: multi-label classification tasks
  - Image tagging, text recognition, named entity recognition, …
- Many multi-label APIs exist
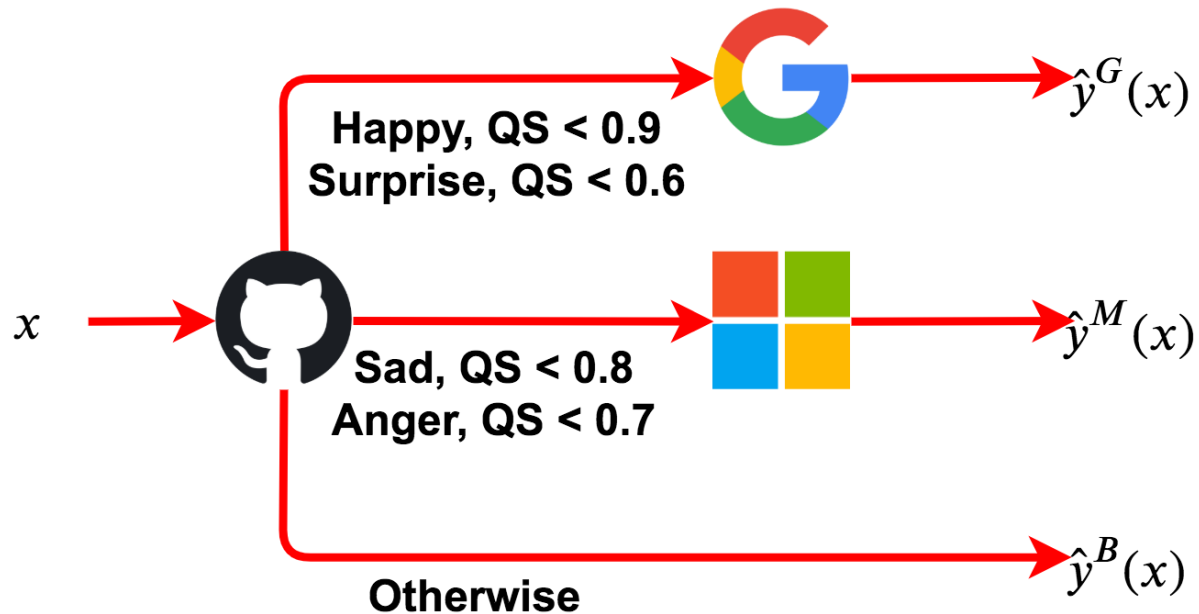- Heterogeneity in those APIs' performance and cost

… …

3

Problem: Which multi-label API to use?

- A common focus: multi-label classification tasks

  - Image tagging, text recognition, named entity recognition, …

- Many multi-label APIs exist

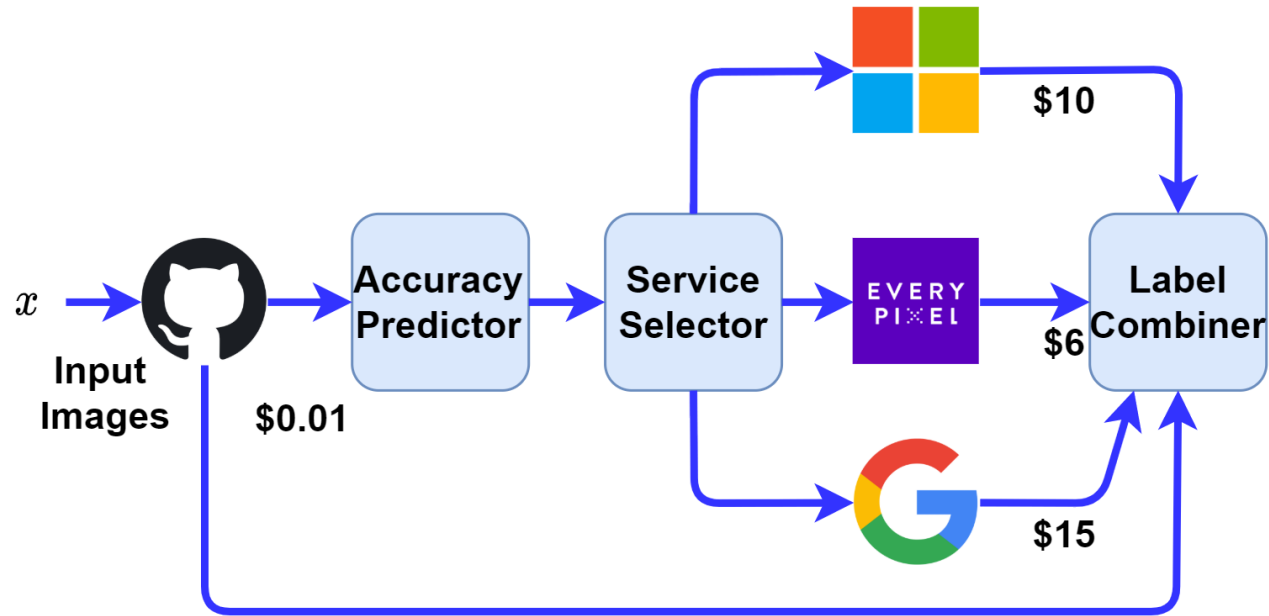- Heterogeneity in those APIs' performance and cost
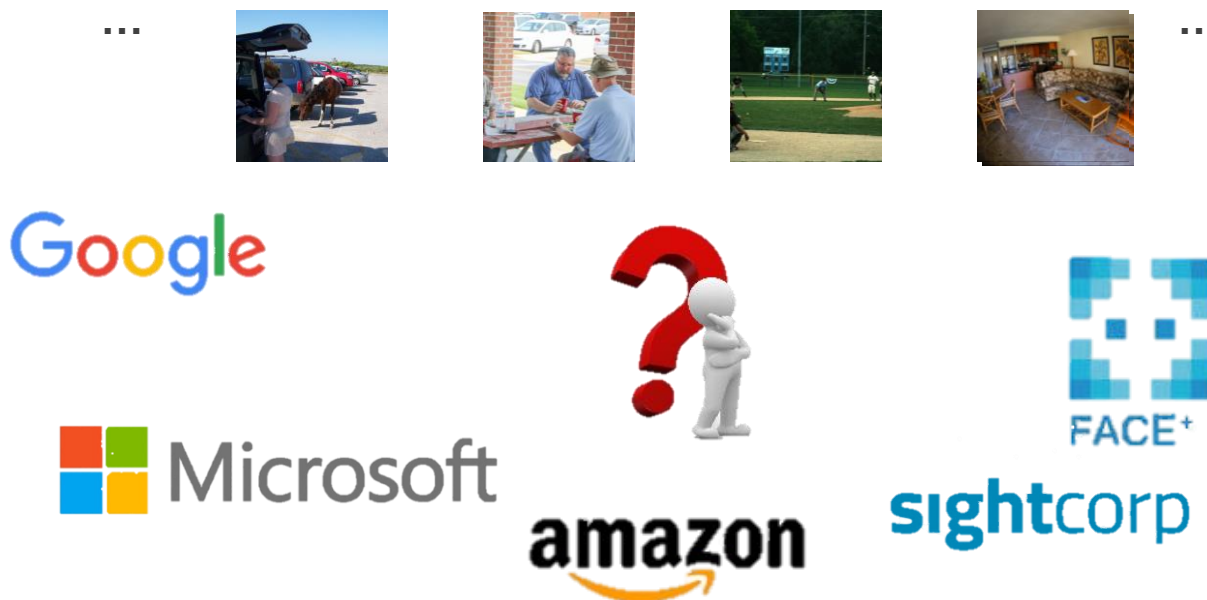
Prior Art for Single-label API Selection

- FrugalML (NeurIPS 2020)
    - Pick API based on Label + Confidence
    - High computational cost for tasks with large # of labels
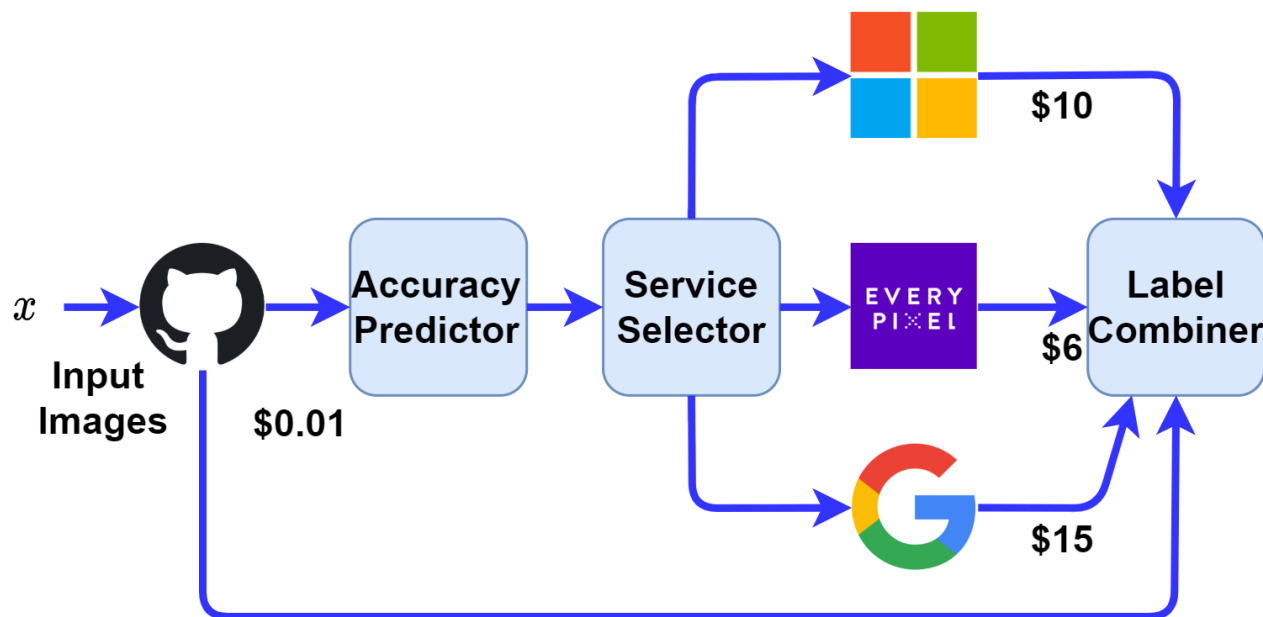   - No prediction fusion for a fixed sample

Our Proposed Solution: FrugalMCT

- A three-stage paradigm

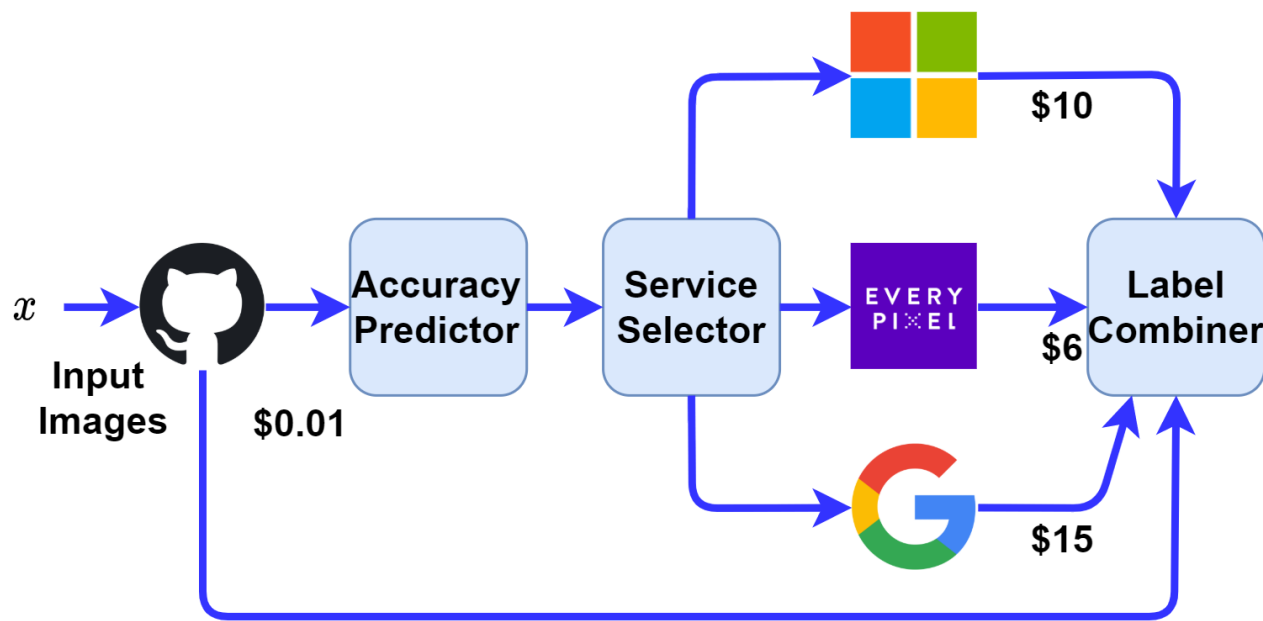# Our Proposed Solution: FrugalMCT

- A three-stage paradigm

✓ **Up to 98% cost savings or 8% better accuracy with same cost across all tasks and datasets evaluated**

# Our Proposed Solution: FrugalMCT

- A three-stage paradigm

✓ **Up to 98% cost savings or 8% better accuracy with same cost across all tasks and datasets evaluated**
✓ **Theoretical analysis on its accuracy and efficiency**
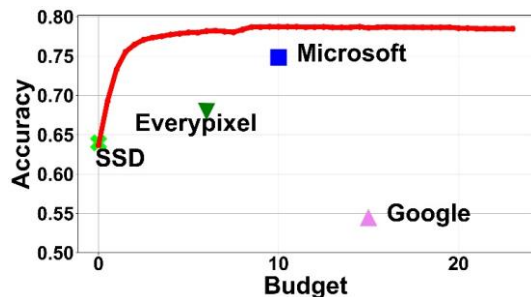
FrugalMCT works well in practice

Task: Image Tagging   Dataset: COCO
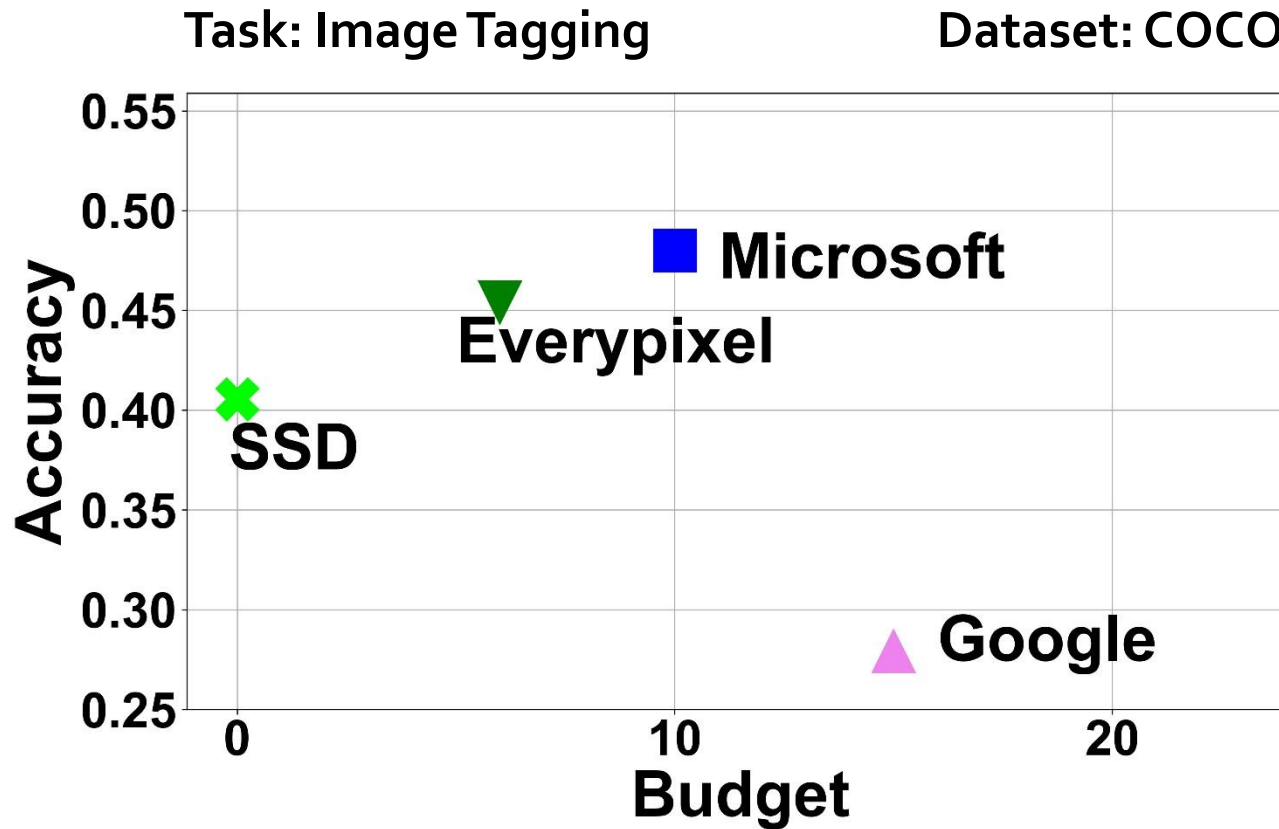
Different APIs' Performance

FrugalMCT works well in practice

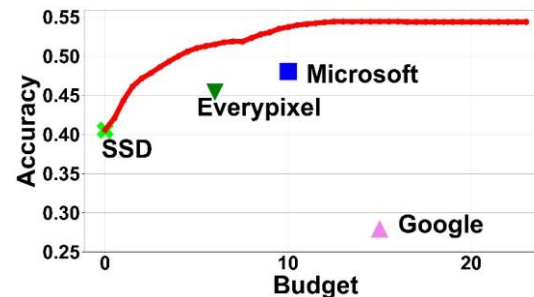Task: Image Tagging   Dataset: COCO
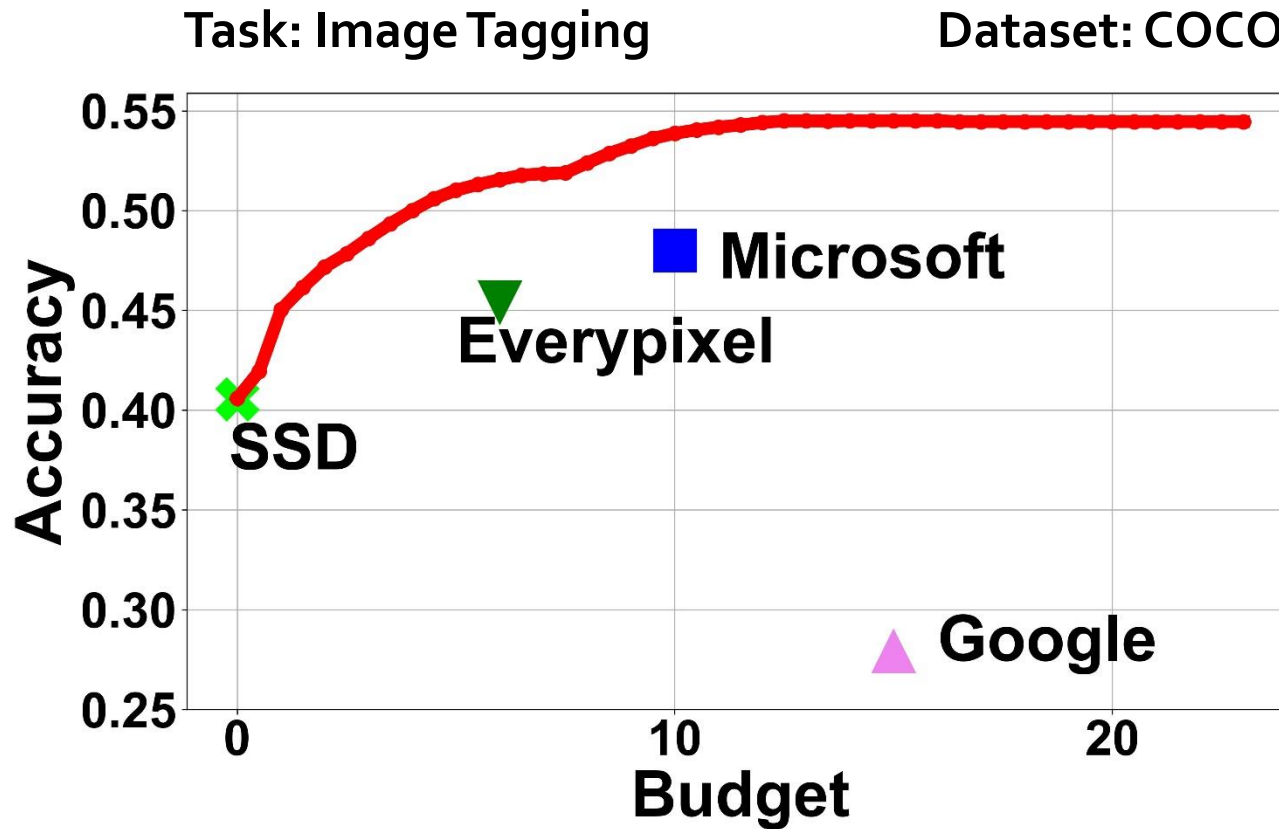
FrugalMCT's performance (red line)
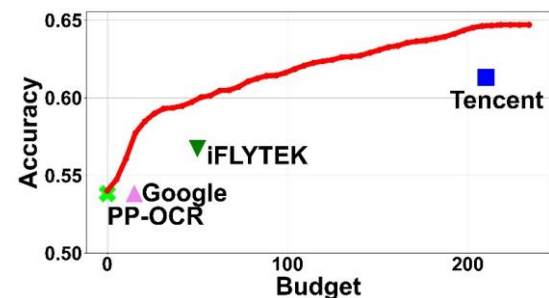
FrugalMCT works well on many datasets
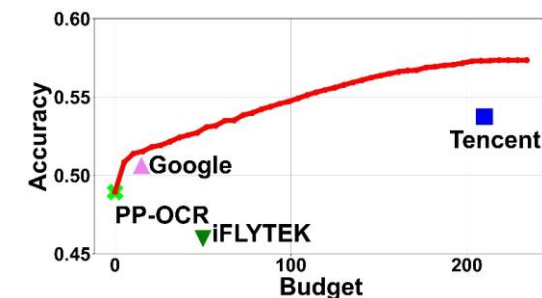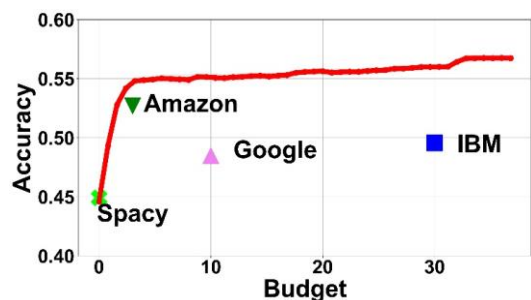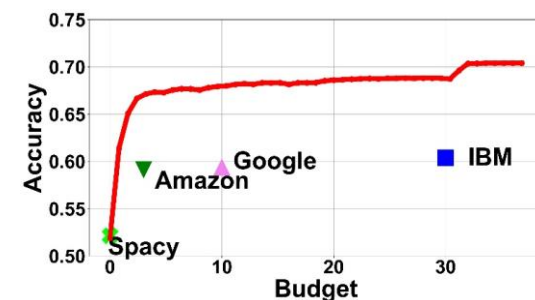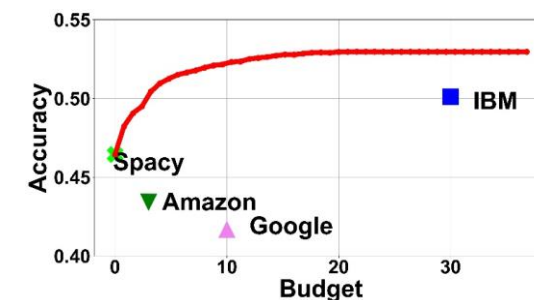

PASCAL
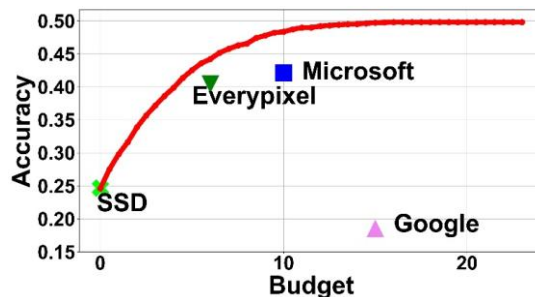
MIR

COCO
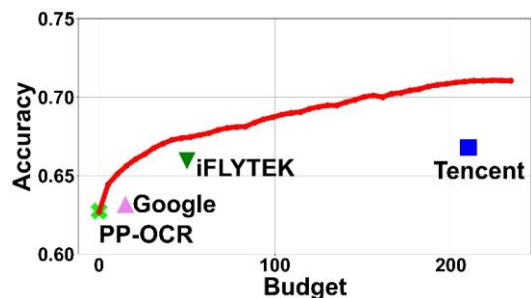
MTWI

ReCTS

LSVT

CONLL

ZHNER

GMB

Takeaways

➢ Question: Best use multi-label ML APIs within a budget

➢ Our solution: FrugalMCT
  ➢ Estimate accuracy, select APIs, and ensemble predictions
  ➢ **Provable** performance and efficiency guarantee
  ➢ Up to **98% cost savings** or **8% better accuracy with same cost**

➢ See the paper for more details