



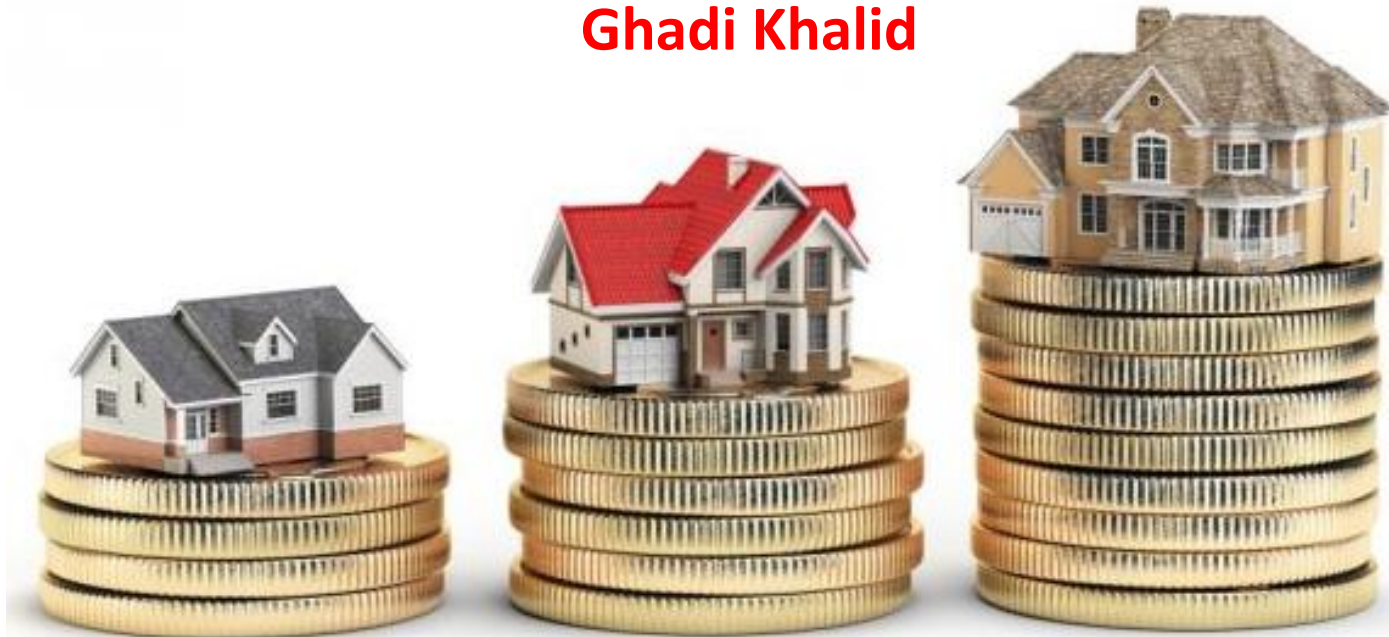
الأكاديمية السعودية الرقمية
SAUDI DIGITAL ACADEMY



Mortgage Prediction

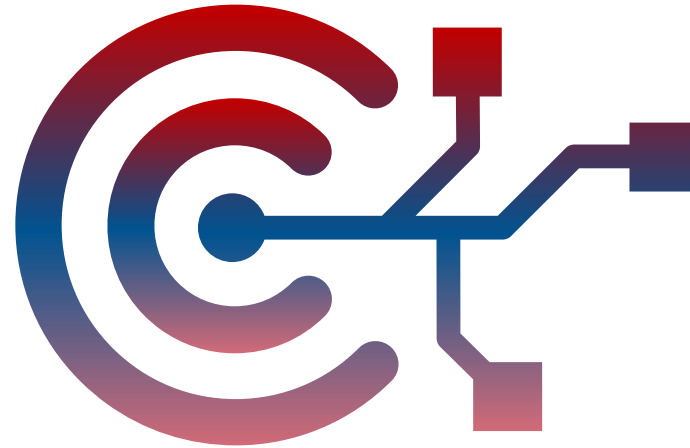
Predict Loan Amount for Customer using RAPIDS Libraries

Ghadi Khalid



AIM OF PROJECT

- ▶ Predict the loan amount using Random Forest model
- ▶ Using RAPIDS libraries cuDF, cuML



DATA SOURCE

Derived From

Sample Data

Dimension

Target

**Fannie Mae's
Single-Family Loan**

2020 3rd

**2,771,993 rows
108 columns**

**Original UPB
dollar amount of
the loan**

1

CLEAN DATA

Add columns names,
Remove Nan,
Fix data format

2

EDA

Explor, Visulize
and Analysis data

3

PREP DATA FOR ML

One-Hot Encoding

4

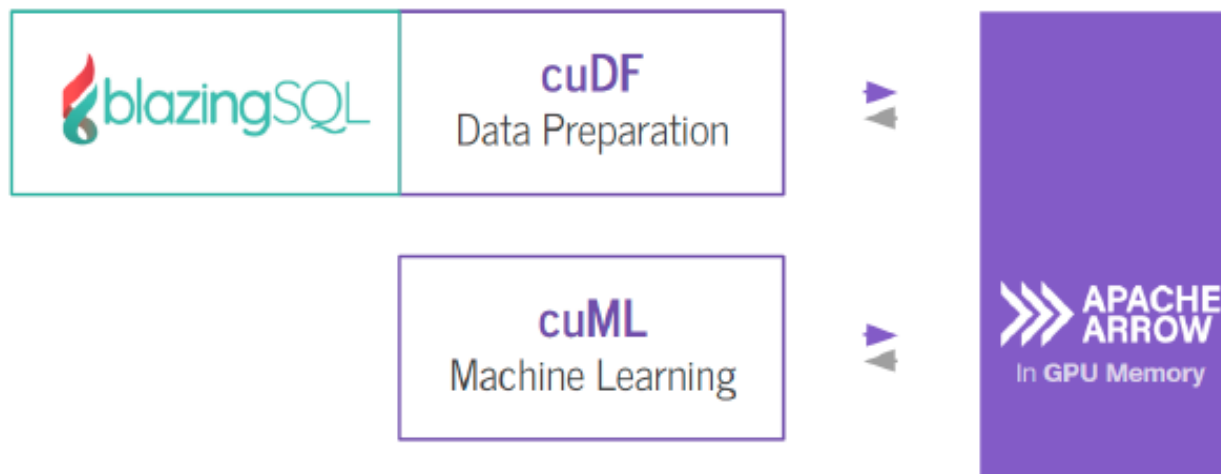
RANDOM FOREST MODEL

Predict loan
amount

CHALLENGES

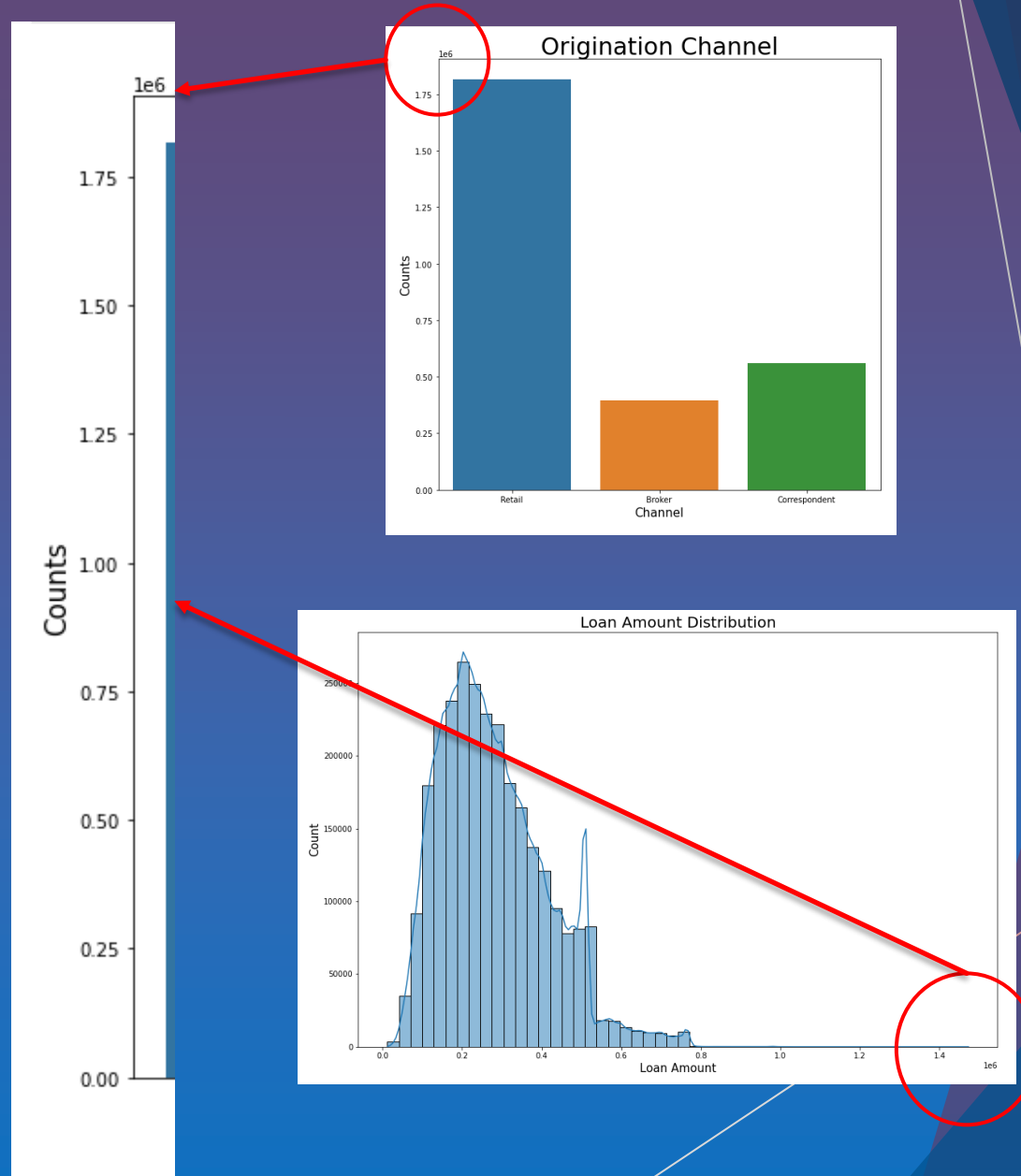
RAPIDS

► Data Size



CHALLENGES

► Scientific Format



CHALLENGES

- Enhance Model
- Larger MSE 66342.09

```
Code
RAPIDS Stable

---> 75         return func(*args, **kwargs)
76         return inner
77

/opt/conda-environments/rapids-stable/lib/python3.8/site-packages/cudf/core/indexing.py in _getitem_tuple_arg(self, arg)
424         df = columns_df._apply_boolean_mask(arg[0])
425     else:
--> 426         df = columns_df._gather(arg[0])
427
428     # Iloc Step 3:

/opt/conda-environments/rapids-stable/lib/python3.8/site-packages/cudf/core/frame.py in _gather(self, gather_map, keep_index)
509         gather_map = gather_map.astype("int32")
510         result = self.__class__._from_table(
--> 511             libcudf.copying.gather(
512                 self, as_column(gather_map), keep_index=keep_index
513             )

cudf/_lib/copying.pyx in cudf._lib.copying.gather()

MemoryError: std::bad_alloc: CUDA error at: /opt/conda-environments/rapids-stable/include/rmm/mr/device/cuda_memory_resource.hpp:69: c
udaErrorMemoryAllocation out of memory
```

```
: # delete unused further data to free up GPU memory
del df_ml , X ,y
```

ANALYSIS INSIGHTS

- ▶ Retail origination channel is of utmost use by the party that delivered the loan to the issuer.
- ▶ The numbers of customers joining Homeready Program in the 3rd Quartier is very low in comparison with customers issuing mortgage but not in Homeready Program.
- ▶ Majority of customers choose to not have Mortgage Insurance.
- ▶ Loan Purpose for the newest loans is either a refinance mortgage or a purchase money mortgage.
- ▶ Most of the customers[borrower or co-borrower] issuing a mortgage are not qualifies as a first-time homebuyer.

MODEL RESULTS

The model predicted the **loan amount**

Performance using **MSE** is **22,663.648**



Off only by

\$22,664



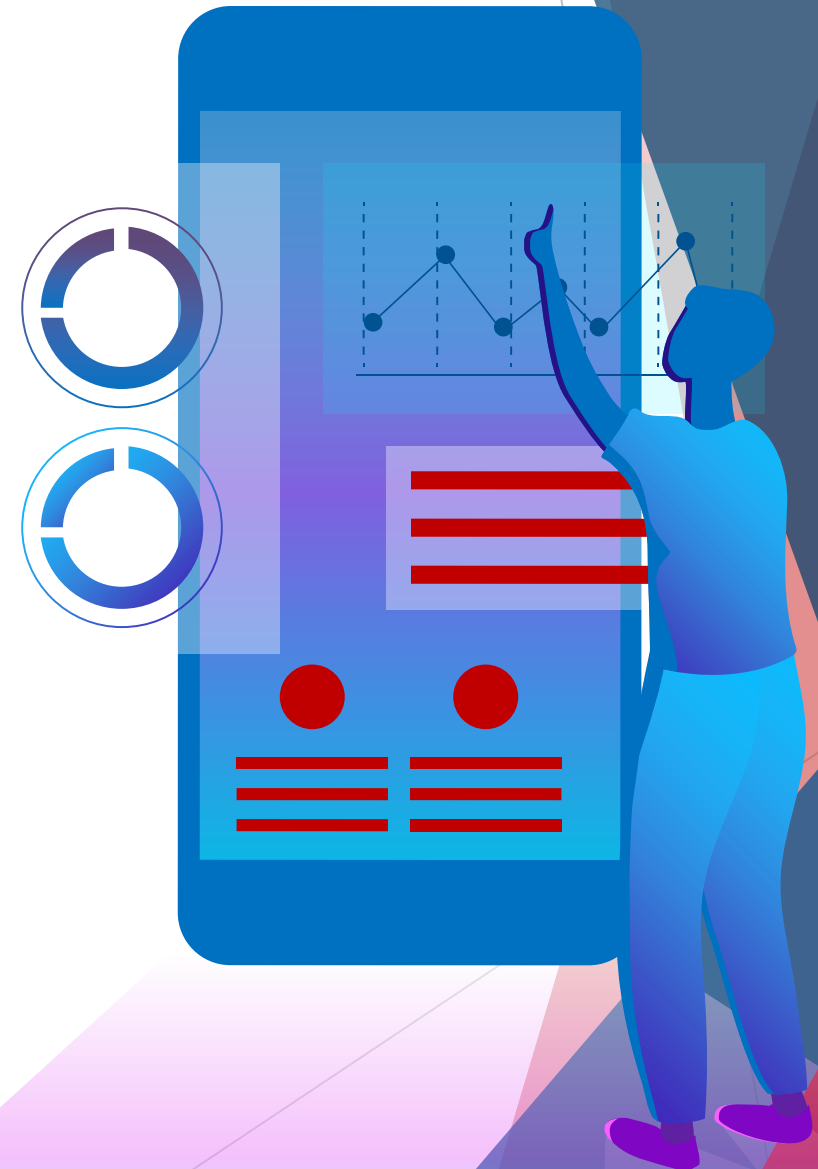
FUTURE DEVELOPMENT



- Reduce predictors
- Interactive visualizations
- Hyperparameters tuning

CONCLUSION

- **Faster loan estimate**
- **Faster customer decision**
- **Improved marketing campaigns**



THANK YOU

