# Which attributes matter the most for loan origination? A neural attention approach



<sup>1</sup>Department of Electrical and Computer Engineering, University of Thessaly, Volos

<sup>2</sup>Department of Electrical Enginnering, Computer Engineering and Informatics, Cyprus University of Technology

### Introduction

 $\sim$ In this work we use an encoder-decoder with attention model to predict loan decisions and provide unique explanations for them.

✓We propose a novel approach for inferring, which attributes matter the most, for making a decision in each specific individual loan case.

- $\sim$ We leverage concepts from neural attention to devise a novel feature wise attention mechanism.
- ✓Our approach offers unique insights into the importance of various features, by producing a decision explanation for each specific loan case.

✓Our novel mechanism generates decisions which are much closer to the decisions generated by human experts, compared to the existent competitors.

 $\sim$ Small introduction to Bahdanau attention:

## Feature Importance Comparisons

Here we present the feature importance results, which the Encoder-Decoder with Attention model produced, and compare them with the results of classical algorithms for feature importance.

Cyprus

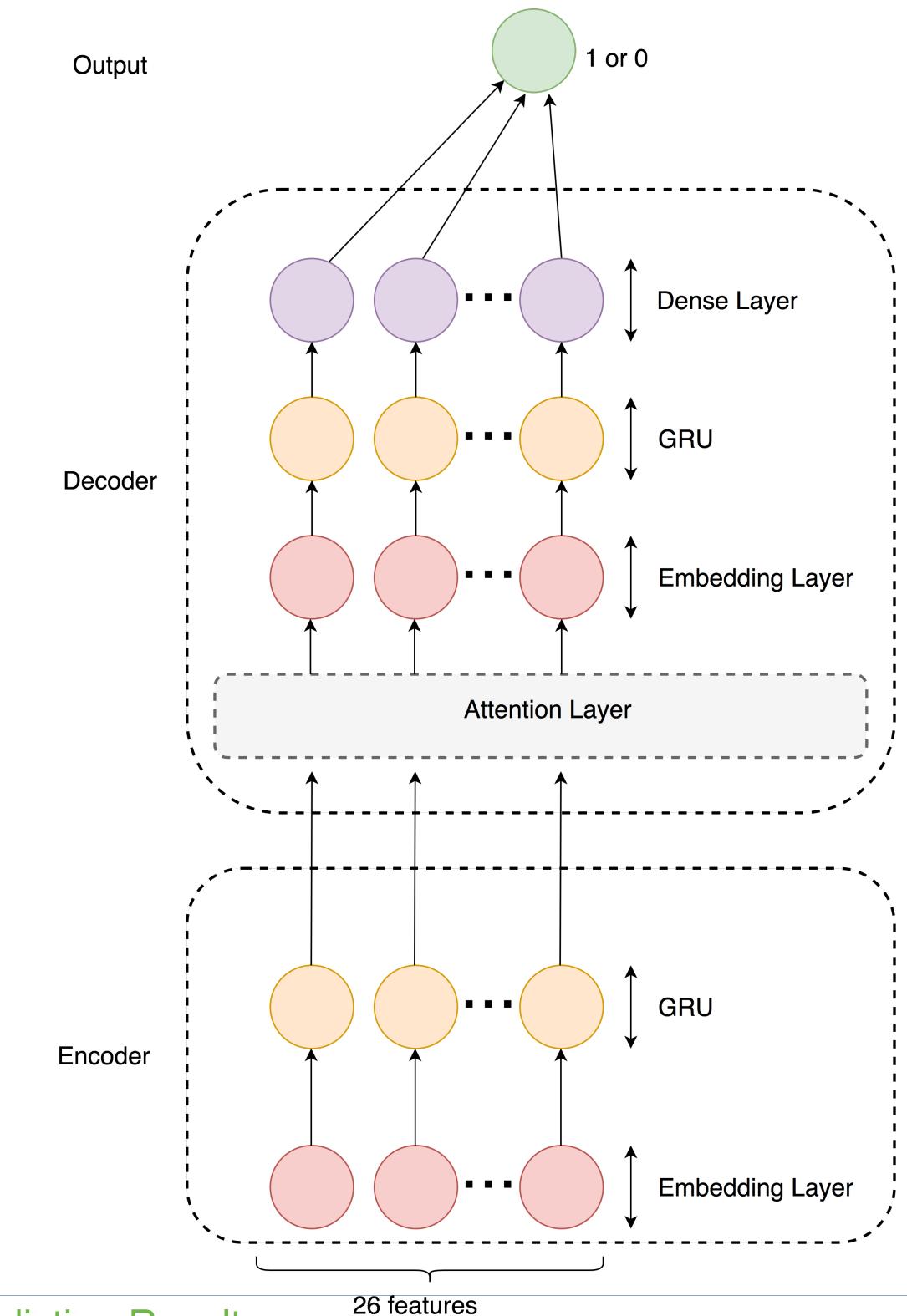
University of

<u> [echnoloay</u>

0	0	0
population -	population -	population -
minority_population -	minority_population -	minority_population -
number_of_owner_occupied_units -	number_of_owner_occupied_units -	number_of_owner_occupied_units -
number_of_1_to_4_family_units -	number_of_1_to_4_family_units -	number_of_1_to_4_family_units
loan_amount_000s -	loan_amount_000s -	loan_amount_000s -
hud_median_family_income -	hud_median_family_income -	hud_median_family_income -
applicant_income_000s -	applicant_income_000s -	applicant_income_000s -
sequence_number-	sequence_number -	sequence_number -
purchaser_type_name -	purchaser_type_name -	purchaser_type_name -
property_type_name -	property_type_name -	property_type_name -
preapproval_name -	preapproval_name -	preapproval_name -
owner_occupancy_name -	owner_occupancy_name -	owner_occupancy_name -
loan_type_name -	loan_type_name -	loan_type_name -
loan_purpose_name -	loan_purpose_name -	loan_purpose_name -
lien_status_name -	lien_status_name -	lien_status_name -
hoepa_status_name -	hoepa_status_name -	hoepa_status_name -
co_applicant_sex_name -	co_applicant_sex_name -	co_applicant_sex_name -
co_applicant_race_name_1 -	co_applicant_race_name_1 -	co_applicant_race_name_1 -
co_applicant_ethnicity_name -	co_applicant_ethnicity_name -	co_applicant_ethnicity_name -
census_tract_number -	census_tract_number -	census_tract_number -
as_of_year-	as_of_year -	as_of_year
applicant_sex_name -	applicant_sex_name -	applicant_sex_name -
applicant_race_name_1 -	applicant_race_name_1 -	applicant_race_name_1 -
applicant_ethnicity_name -	applicant_ethnicity_name -	applicant_ethnicity_name -
agency_name -	agency_name -	agency_name -
Decision Tree 84.7%	Extra Trees 889	LASSO
Decision Tree <b>84.7%</b>	Extra Trees 889	LASSO
Decision Tree 84.7%	Extra Trees 889	b LASSO
o population - minority_population -	0 population - minority_population -	population - minority_population -
o population - minority_population - number_of_owner_occupied_units -	0 - population - minority_population - number_of_owner_occupied_units	population - minority_population - number_of_owner_occupied_units -
o population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units -	0 population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s -	0 population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number -	population - population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name -	population - population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name -	population - population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name -
0 population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - owner_occupancy_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - owner_occupancy_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - owner_occupancy_name -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - owner_occupancy_name - loan_type_name -	population = minority_population = number_of_owner_occupied_units = number_of_1_to_4_family_units = loan_amount_000s = hud_median_family_income = applicant_income_000s = sequence_number = purchaser_type_name = property_type_name = preapproval_name = owner_occupancy_name = loan_type_name =	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name -
population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_purpose_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_purpose_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name -
0 population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - preapproval_name - loan_type_name - loan_type_name - loan_purpose_name - lien_status_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - preapproval_name - loan_type_name - loan_type_name - loan_purpose_name - lien_status_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_purpose_name -
0 population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - loan_type_name - loan_type_name - loan_purpose_name - lien_status_name - hoepa_status_name -	o population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - lien_status_name - hoepa_status_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name -
population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - loan_type_name - loan_type_name - loan_purpose_name - lien_status_name - hoepa_status_name - co_applicant_sex_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_purpose_name - lien_status_name - hoepa_status_name - co_applicant_sex_name -	population - population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - property_type_name - loan_type_name - loan_
population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - loan_type_name - loan_type_name - lien_status_name - hoepa_status_name - co_applicant_race_name_1 -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - lien_status_name - hoepa_status_name - co_applicant_sex_name - co_applicant_race_name_1 -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_purpose_name - lien_status_name - hoepa_status_name - co_applicant_sex_name -
population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_purpose_name - lien_status_name - hoepa_status_name - co_applicant_sex_name - co_applicant_race_name_1 - co_applicant_ethnicity_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_purpose_name - lien_status_name - hoepa_status_name - co_applicant_sex_name - co_applicant_race_name_1 - co_applicant_ethnicity_name -	population = population = number_of_owner_occupied_units = number_of_1_to_4_family_units = loan_amount_000s = hud_median_family_income = applicant_income_000s = sequence_number = purchaser_type_name = property_type_name = property_type_name = loan_type_name = loan_type_name = lien_status_name = hoepa_status_name = co_applicant_sex_name = co_applicant_race_name_1 =
population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - gequence_number - purchaser_type_name - property_type_name - loan_type_name - loan_type_name - loan_purpose_name - lien_status_name - hoepa_status_name - co_applicant_sex_name - co_applicant_race_name_1 - co_applicant_race_name_1	population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - co_applicant_sex_name - co_applicant_race_name_1 - co_applicant_race_name_1	population = population = minority_population = number_of_owner_occupied_units = number_of_1_to_4_family_units = loan_amount_000s = hud_median_family_income = applicant_income_000s = sequence_number = purchaser_type_name = property_type_name = property_type_name = loan_type_name = loan_
<pre>population -     minority_population -     minority_population -     number_of_owner_occupied_units     number_of_1_to_4_family_units     loan_amount_000s -     loan_amount_000s -     hud_median_family_income -     applicant_income_000s -     sequence_number -     purchaser_type_name -     property_type_name -     property_type_name -     loan_type_name -     loan_purpose_name -     lien_status_name -     hoepa_status_name -     co_applicant_race_name_1     co_applicant_race_name_1     co_applicant_ethnicity_name -     co_applicant_ethnici</pre>	population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_type_name - lien_status_name - hoepa_status_name - co_applicant_sex_name - co_applicant_race_name_1 - co_applicant_race_name_1 - co_applicant_race_name_1 - co_applicant_ethnicity_name - census_tract_number - as_of_year -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - property_type_name - loan_type_name - loan_type
population - minority_population - number_of_owner_occupied_units number_of_1_to_4_family_units loan_amount_000s - hud_median_family_income applicant_income_000s sequence_number purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - co_applicant_sex_name - co_applicant_race_name_1 co_applicant_race_name_1 co_applicant_race_name_1 co_applicant_race_name_1 co_applicant_ethnicity_name - as_of_year - applicant_sex_name -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - loan_type_name - co_applicant_sex_name - co_applicant_sex_name - co_applicant_race_name_1 co_applicant_ethnicity_name - census_tract_number - as_of_year - applicant_sex_name -	<pre></pre>
<pre>population -     minority_population -     minority_population -     number_of_owner_occupied_units     number_of_1_to_4_family_units     loan_amount_000s -     loan_amount_000s -     hud_median_family_income -     applicant_income_000s -     sequence_number -     purchaser_type_name -     property_type_name -     property_type_name -     loan_type_name -     loan_purpose_name -     lien_status_name -     hoepa_status_name -     co_applicant_race_name_1     co_applicant_race_name_1     co_applicant_ethnicity_name -     co_applicant_ethnici</pre>	population - minority_population - number_of_owner_occupied_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - preapproval_name - loan_type_name - loan_type_name - loan_type_name - lien_status_name - hoepa_status_name - co_applicant_sex_name - co_applicant_race_name_1 - co_applicant_race_name_1 - co_applicant_race_name_1 - co_applicant_ethnicity_name - census_tract_number - as_of_year -	population - minority_population - number_of_owner_occupied_units - number_of_1_to_4_family_units - loan_amount_000s - hud_median_family_income - applicant_income_000s - sequence_number - purchaser_type_name - property_type_name - property_type_name - loan_type_name - loan_type

## Methodology

- In our work, we treat every observation(vector) loan data as a sequence of tokens, with each feature being treated as a token by the model.
- ✓We make the assumption that the data are not sequential, and do not have any dependencies.
  ✓This is a slight violation to the model.
- →By making this assumption we explicitly consider sequential dynamics. The form of dynamics of dependencies is adaptive due to the attention mechanism.
- $\sim$ So this slight violation that exists, is not detrimental due to the attention mechanism.



XGBoost 89%	Our approach (	Our approach 1
0		0 0
population -	population -	population -
minority_population -	minority_population -	minority_population -
number_of_owner_occupied_units -	number_of_owner_occupied_units -	number_of_owner_occupied_units
number_of_1_to_4_family_units -	number_of_1_to_4_family_units -	number_of_1_to_4_family_units -
loan_amount_000s -	loan_amount_000s -	loan_amount_000s -
hud_median_family_income -	hud_median_family_income -	hud_median_family_income -
applicant_income_000s -	applicant_income_000s -	applicant_income_000s -
sequence_number -	sequence_number -	sequence_number -
purchaser_type_name -	purchaser_type_name -	purchaser_type_name -
property_type_name -	property_type_name -	property_type_name -
preapproval_name -	preapproval_name -	preapproval_name -
owner_occupancy_name -	owner_occupancy_name -	owner_occupancy_name -
loan_type_name -	loan_type_name -	loan_type_name -
loan_purpose_name -	loan_purpose_name -	loan_purpose_name -
lien_status_name -	lien_status_name -	lien_status_name -
hoepa_status_name -	hoepa_status_name -	hoepa_status_name -
co_applicant_sex_name	co_applicant_sex_name -	co_applicant_sex_name -
co applicant race name 1	co_applicant_race_name_1 -	co_applicant_race_name_1 -
co applicant ethnicity name	co_applicant_ethnicity_name -	co_applicant_ethnicity_name -
census_tract_number -	census_tract_number -	census_tract_number-
as_of_year	as_of_year-	as_of_year-
applicant_sex_name -	applicant_sex_name -	applicant_sex_name -
applicant_race_name_1	applicant_race_name_1 -	applicant_race_name_1 -
applicant_ethnicity_name -	applicant_ethnicity_name -	applicant_ethnicity_name -
agency_name	agency_name -	agency_name -
Our approach 2	Our approach 3	B Our approach 4

agency\_name -

agency\_name -



agency\_name -

 $\sim$ The Encoder-Decoder with attention model has clearly other utilities apart from translating

#### **Prediction Results**

Lo louidio

In order to prove the superiority of our model we compare its prediction results with the results of a simple neural network. The simple neural network achieved a prediction score of 73.2%, while our model achieved a prediction score of 89.6%.

Layer (type)	Output Shape	Param #	
dense_4 (Dense)	(None, 100)	2700	
dropout_3 (Dropout)	(None, 100)	0	

dense_5 (Dense)	(None, 100)	10100
dropout_4 (Dropout)	(None, 100)	0
dense_6 (Dense)	(None, 1)	101
Total params: 12,901 Trainable params: 12,901 Non-trainable params: 0		

	Yes	No		Yes
Yes	0	1914	Yes	1307
No	0	5243	No	179

Confusion matrix of the simple Neural Network

Confusion matrix of our model

No

559

5059

- sentences in other languages. It has the ability to show unique feature importance for every observation.
- ✓We have successfully achieved, to explain bank loan decisions to the applicants, by giving them unique explanations regarding their application.
- Through data analysis we have seen that the algorithm is not discriminating in the rejection cases. Features like race and ethnicity don not play a major role in a rejected case.

#### Future Work

- $\sim$ We will try to implement BERT on the data, regarding its unique
- $\sim$ Change the order of the feature and observe any changes on the accuracy.
- $\sim$ Train the model with the whole dataset in order to improve the accuracy score.
- $\sim$ Try to implement the idea to another dataset.