

A decorative network diagram in the top-left corner, featuring a cluster of interconnected nodes. Some nodes are solid blue circles, while others are white circles with blue outlines. They are connected by thin grey lines.

# Real Estate Investment Application

Scott Rosengrants  
DEN - DSI - 10 Capstone Project

A decorative network diagram in the bottom-right corner, similar to the one in the top-left, with a cluster of interconnected nodes. Some nodes are solid blue circles, while others are white circles with blue outlines. They are connected by thin grey lines.

## What RE Investors Currently Do

1. Manually search and calculate rental properties
2. Search inconsistently for fair market rents
3. Call a lender each time a property is considered





## Project Objective - Problem Statement

1. Given a property address can relevant rental amounts be found for the property? Can standard real estate metrics be automated?
2. According to the location of the property and user provided financial information can a model be developed for predicting the approval of a potential mortgage loan?

## Start with the Data

1. Property Records

- ATTOM Data Solutions API



2. Fair Market Rental Information

- County Level Fair Market Rents FY 2020

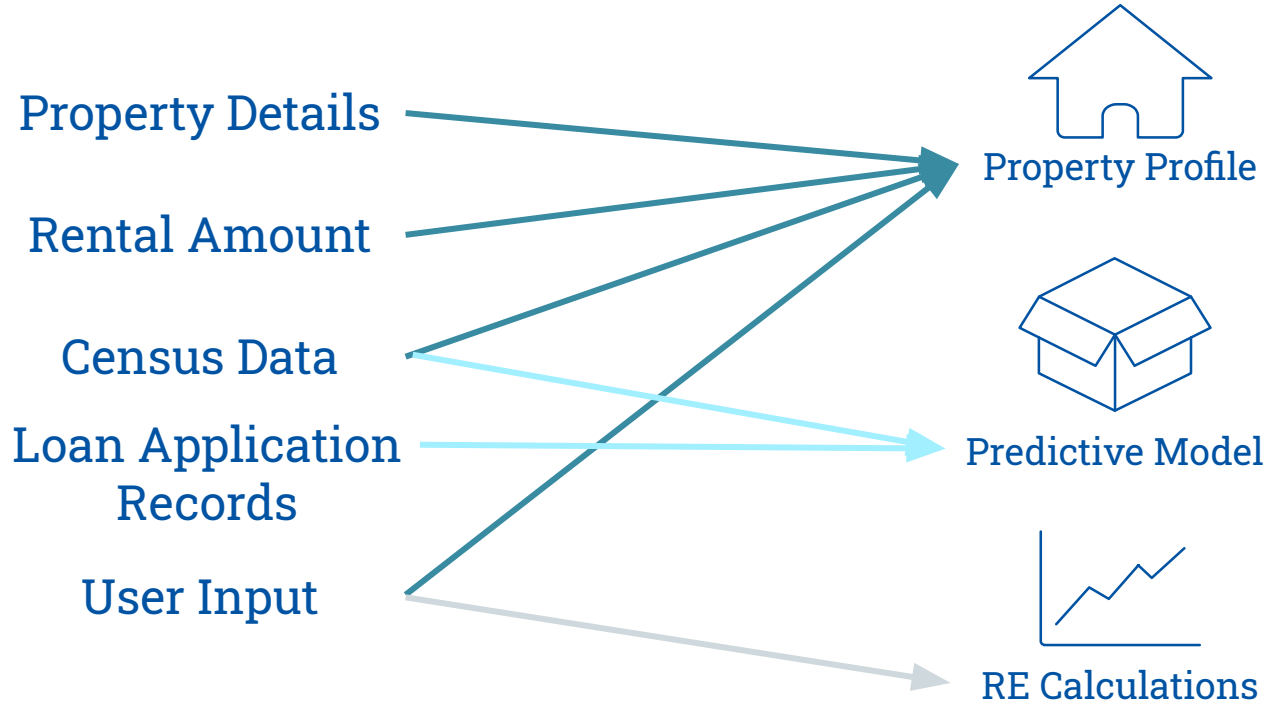


3. Mortgage Records

- 2018 Loan/Applications Records



## Data Flow





## Elements in the Property Profile

- ⦿ Location and Census demographic information
- ⦿ Property Attributes  
(Bedrooms, Bathrooms, Square footage)
- ⦿ Fair market rent based on County
- ⦿ User Inputs  
(Price, income, type of loan)



## Predictive Model

## Model Prediction Classes

### Two Classification

Approved

Denied

### Ten Classification

Approved

Debt-to-income ratio

Employment history

Denied

Credit history

Collateral

Insufficient cash  
(downpayment, closing costs)

Unverifiable information

Credit application incomplete

Mortgage insurance denied

Other



# Predictive Model

## Models Attempted

Model Type	Accuracy
Logistic Regression (10-Class)	<b>91.276%</b>
Logistic Regression (2-Class)	<b>91.276%</b>
FF Neural Network (10-Class)	<b>92.11%</b>
FF Neural Network (2-Class)	<b>94.15%</b>





## Predictive Model

### Inputs to the Model

- Loan type information
- Census area demographics
- User's income
- Property value
- Location
- 72 elements in total

### Deep Feed Forward (DFF)



### Targets

- Approved for mortgage
- Denied for mortgage

Model Accuracy

**94.15%**



## RE Calculations

### **Cash Flow Calculation**

(monthly or annual)

Income (Rent)

- Vacancy Rate
- Expense Reserves
- Taxes
- Insurance
- Financing Expense

---

Cash Flow



## RE Calculations

### **Cash on Cash Return (COC) Calculation**

$$\frac{\text{Annual Cash Flow}}{\text{Total Cash Invested}} \times 100$$



## RE Calculations

### **Capitalization Rate (Cap Rate) Calculation**

(no financing costs)

$$\frac{\text{Annual Cash Flow}}{\text{Price of the Property}} \times 100$$

**Property Report for: 33 S COLORADO AVE, INDIANAPOLIS, IN 46201**

Monthly Income: \$1534

Vacancy Expense (5%): \$76.7

Expense Reserves (10%): \$153.4

Monthly Expenses(Trash & Water): \$159.8

Monthly Taxes: \$82.83

Monthly Insurance: \$82.83

Monthly Financing Expense: \$164.71

Monthly Cashflow: \$823.52

Annual Cashflow: \$9882.28

Cash-on-Cash Return: 28.64%

Capitalization Rate: 17.19%

ATTOM AVM Poor-Quality Range:

\$38400.0 - \$74100.0

ATTOM AVM Good-Quality:

\$32000.0 - \$78000.0

ATTOM AVM Excellent-Quality:

\$32800.0 - \$79950.0

You are likely to qualify for a mortgage loan on this particular property

This application is for educational purposes only. No warranties are made on the completeness, reliability, and accuracy of this information

# Putting It All Together

Flask Application Example



## Large Limitations


The rental data is only based on County, this limits the accuracy of this information.

The model is a “black-box” model resulting more accurate predictions but unfortunately little to no insight to the user as to why they may be denied for a loan.

The data used to train the model was limited and may not be indicative to actual approval odds.

1. Debt - to - Income ratio
2. Credit Score

The model likely picked up on a trend that is unseen in the data. Limiting the ability to replicate results.



## Improvements & Additional Work

- ◎ Credit Score and Debt-to-Income - higher accuracy and usability of model predictions
- ◎ Neighborhood Specific Rental Figures - multiple rental figures per county
- ◎ Multiple Unit Analysis - multifamily properties
- ◎ A pro-active application - emails based on user settings

# Sources

ATTOM Data Solutions

<https://api.developer.attomdata.com/>

Rental Data

[https://www.huduser.gov/portal/datasets/fmr.html#2020\\_data](https://www.huduser.gov/portal/datasets/fmr.html#2020_data)

Mortgage Data

<https://ffiec.cfpb.gov/data-publication/dynamic-national-loan-level-dataset/2018>

Neural Network Image

<https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464>

Flask Framework

<https://coreyms.com/>






# Thanks!

## Any questions?

You can find me at:  
[linkedin.com/in/scott-rosengrants](https://www.linkedin.com/in/scott-rosengrants)



This template is free to use under [Creative Commons Attribution license](#).  
You can keep the Credits slide or mention SlidesCarnival and other  
resources used in a slide footer.