

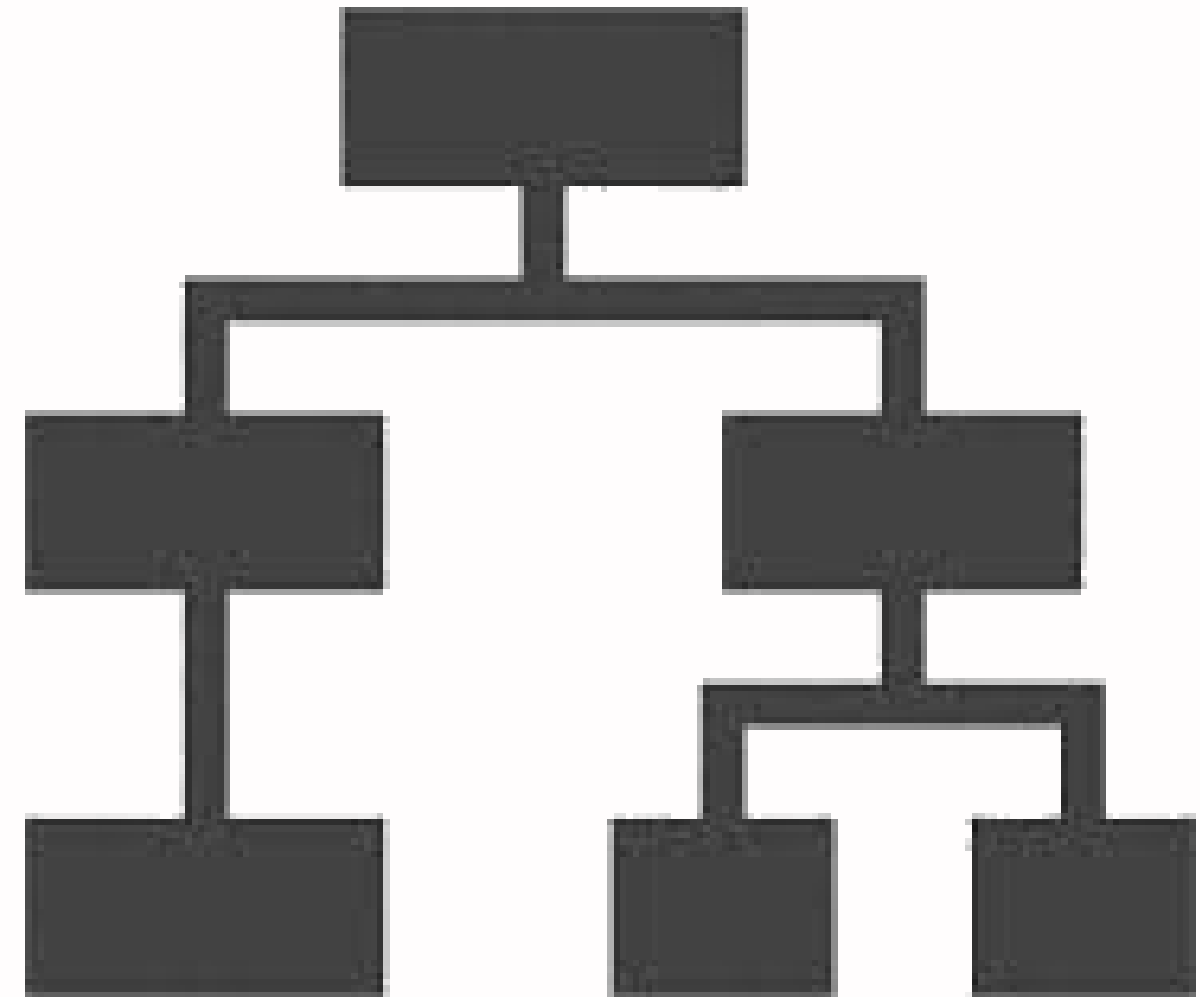
# DOCUMENT CLASSIFICATION

Loan Servicers Receive Millions of Forms

Forms need to be hand coded for routing

Can we automate this document routing  
process?

PRESENTED BY PATRICK BROWN



# DOCUMENTS



Direct Consolidation Loan Application and Promissory Note

William D. Ford Federal Direct Loan Program

OMB No. 1845-0187

Form Approved

Exp. Date 09/30/2023

**WARNING:** Any person who knowingly makes a false statement or misrepresentation on this form or any accompanying document is subject to penalties that may include fines, imprisonment, or both, under the U.S. Criminal Code and 20 U.S.C. 1097.

**BEFORE YOU BEGIN**

Read the Instructions for Completing the Direct Consolidation Loan Application and Promissory Note.

**NOTE: PAGES 1 THROUGH 5 MUST BE SUBMITTED FOR YOUR LOAN APPLICATION TO BE PROCESSED.**

**BORROWER INFORMATION**

1. Last Name: \_\_\_\_\_ First Name: \_\_\_\_\_ Middle Initial: \_\_\_\_\_

2. Former Name(s): \_\_\_\_\_

3. Social Security Number: \_\_\_\_\_

4. Date of Birth: \_\_\_\_\_

5. Permanent Address (Street, City, State, Zip Code) (If P.O. box or general delivery, see Instructions): \_\_\_\_\_

6. Area Code/Telephone Number: \_\_\_\_\_

7. Email Address (optional): \_\_\_\_\_

8. Driver's License State and Number: \_\_\_\_\_

9. Employer's Name and Address (Street, City, State, Zip Code): \_\_\_\_\_

10. Work Area Code/Telephone Number: \_\_\_\_\_

**REFERENCE INFORMATION**

List two persons with different U.S. addresses who do not live with you and who have known you for at least three years.

11. Last Name: \_\_\_\_\_ First Name: \_\_\_\_\_ Middle Initial: \_\_\_\_\_

Permanent Address (Street, City, State, Zip Code): \_\_\_\_\_

Email Address (optional): \_\_\_\_\_

Area Code/Telephone Number: \_\_\_\_\_

Relationship to You: \_\_\_\_\_

12. Last Name: \_\_\_\_\_ First Name: \_\_\_\_\_ Middle Initial: \_\_\_\_\_

Permanent Address (Street, City, State, Zip Code): \_\_\_\_\_

Email Address (optional): \_\_\_\_\_


Area Code/Telephone Number: \_\_\_\_\_

Relationship to You: \_\_\_\_\_

Submit PAGES 1 THROUGH 5  
Page 1 of 16

12/2019

DLC



INCOME-DRIVEN REPAYMENT (IDR) PLAN REQUEST

For the Revised Pay As You Earn (REPAYE), Pay As You Earn (PAYE),  
Income-Based Repayment (IBR), and Income-Contingent Repayment (ICR)  
plans under the William D. Ford Federal Direct Loan (Direct Loan) Program  
and Federal Family Education Loan (FFEL) Programs

OMB No. 1845-0102

Form Approved

Expiration Date:  
8/31/2021

**WARNING:** Any person who knowingly makes a false statement or misrepresentation on this form or on any accompanying document is subject to penalties that may include fines, imprisonment, or both, under the U.S. Criminal Code and 20 U.S.C. 1097.

**SECTION 1: BORROWER INFORMATION**

Please enter or correct the following information.

☐ Check this box if any of your information has changed.

SSN \_\_\_\_\_

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip Code \_\_\_\_\_

Telephone - Primary \_\_\_\_\_

Telephone - Alternate \_\_\_\_\_

Email (Optional) \_\_\_\_\_

**SECTION 2: REPAYMENT PLAN OR RECERTIFICATION REQUEST**

It's faster and easier to complete this form online at [StudentAid.gov](http://StudentAid.gov). You can learn more at [StudentAid.gov/IDR](http://StudentAid.gov/IDR) and by reading Sections 9 and 10. It's simple to get repayment estimates at [StudentAid.gov/repayment-estimator](http://StudentAid.gov/repayment-estimator). If you need help with this form, contact your loan holder or service for free assistance. You can find out who your loan holder or service is at [StudentAid.gov/login](http://StudentAid.gov/login). You may have to pay income tax on any loan amount forgiven under an income-driven plan.

1. Select the reason you are submitting this form (Check only one):

☐ I want to enter an income-driven plan - Continue to Item 2.

☐ I am submitting documentation for the annual recertification of my income-driven payment - Skip to Item 3.

☐ I am submitting documentation early to have my income-driven payment recalculated immediately - Skip to Item 3.

☐ I want to change to a different income-driven plan - Continue to Item 2.

2. Choose a plan and then continue to Item 3.

☐ (Recommended) I want the income-driven repayment plan with the lowest monthly payment.

☐ REPAYE ☐ IBR ☐ PAYE ☐ ICR

3. Do you have multiple loan holders or servicers?

☐ Yes - Submit a request to each holder or servicer. Continue to Item 4.

☐ No - Continue to Item 4.

4. Are you currently in deferment or forbearance? After answering, continue to Item 5.

☐ No.

☐ Yes, but I want to start making payments under my plan immediately.

☐ Yes, and I do not want to start repaying my loans until the deferment or forbearance ends.

Note: If you have FFEL Program loans, they are only eligible for IBR. However, you can consolidate your loans at [StudentAid.gov](http://StudentAid.gov) to access more beneficial income-driven repayment plans.

**SECTION 3: FAMILY SIZE INFORMATION**


5. How many children, including unborn children, are in your family and receive more than half of their support from you? \_\_\_\_\_

6. How many other people, excluding your spouse and children, live with you and receive more than half of their support from you? \_\_\_\_\_

Note: A definition of "family size" is provided in Section 9. Do not enter a value for you or your spouse. Those values are automatically included in your family size, if appropriate.

Page 1 of 10

IDR



TEACHER LOAN FORGIVENESS FORBEARANCE REQUEST

William D. Ford Federal Direct Loan (Direct Loan) Program  
Federal Family Education Loan (FFEL) Program

OMB No. 1845-0059

Form Approved

Exp. Date 09/30/2023

**WARNING:** Any person who knowingly makes a false statement or misrepresentation on this form or on any accompanying document is subject to penalties that may include fines, imprisonment, or both, under the U.S. Criminal Code and 20 U.S.C. 1097.

**SECTION 1: BORROWER IDENTIFICATION**

Please enter or correct the following information.

☐ Check this box if any of your information has changed.

SSN \_\_\_\_\_

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip Code \_\_\_\_\_

Telephone - Primary \_\_\_\_\_

Telephone - Alternate \_\_\_\_\_

Email (Optional) \_\_\_\_\_

**SECTION 2: LOAN FORGIVENESS FORBEARANCE REQUEST**

**READ SECTIONS 5 - 8 BEFORE COMPLETING THE FORM.**

I request forbearance of payments on my eligible Direct Loan and/or FFEL program loans while I am performing qualifying teaching service as a full-time teacher at an elementary or secondary school or for an educational service agency. During the period that qualifies me for this forbearance, I am teaching (check one):

☐ at an eligible elementary school ☐ at an eligible secondary school ☐ for an eligible educational service agency

AND I am (check all that apply):

☐ A highly qualified full-time special education teacher for elementary school children with disabilities (forgiveness of up to \$17,500).

☐ A highly qualified full-time special education teacher for secondary school children with disabilities (forgiveness of up to \$17,500).

☐ A highly qualified full-time mathematics teacher for secondary school students (forgiveness of up to \$17,500).


☐ A highly qualified full-time science teacher for secondary school students (forgiveness of up to \$17,500).

☐ A highly qualified full-time secondary education teacher, or (only if my teaching service began before 10/30/2004) a full-time secondary education teacher in a subject area relevant to my academic major (forgiveness of up to \$5,000).

A highly qualified full-time elementary education teacher, or (only if my teaching service began before 10/30/2004) a full-time elementary education teacher and I demonstrated knowledge and teaching skills in reading, writing, mathematics, and other areas of the elementary school curriculum (forgiveness of up to \$5,000).

Page 1 of 7

TLF



ECONOMIC HARDSHIP DEFERMENT REQUEST

William D. Ford Federal Direct Loan (Direct Loan) Program / Federal Family Education Loan (FFEL) Program / Federal Perkins Loan (Perkins Loan) Program

OMB No. 1845-0011

Form Approved

Exp. Date 8/31/2021

**WARNING:** Any person who knowingly makes a false statement or misrepresentation on this form or on any accompanying document is subject to penalties that may include fines, imprisonment, or both, under the U.S. Criminal Code and 20 U.S.C. 1097.

**SECTION 1: BORROWER INFORMATION**

Please enter or correct the following information.

☐ Check this box if any of your information has changed.

SSN \_\_\_\_\_

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip Code \_\_\_\_\_

Telephone - Primary \_\_\_\_\_

Telephone - Alternate \_\_\_\_\_

Email (Optional) \_\_\_\_\_

**SECTION 2: BORROWER DETERMINATION OF DEFERMENT ELIGIBILITY**

**Maximum cumulative eligibility is 36 months per loan program.** Instead of deferment, consider a repayment plan that bases your monthly payment amount on your income. Visit [StudentAid.gov/IDR](http://StudentAid.gov/IDR) for more information.

1. Have you received this type of deferment on a loan made under another federal student loan program for the same period of time that you want this deferment? For example, check "yes" if you want this deferment on your Direct Loans and you are on this deferment on your FFEL Program loans.

☐ Yes - Attach documentation of the deferment. Skip to Section 3.

☐ No - Continue to Item 2.

2. Have you received or are you receiving payments under a federal or state public assistance program for the same period of time that you want this deferment? Qualifying programs include: Temporary Assistance for Needy Families (TANF), Supplemental Security Income (SSI), Supplemental Nutrition Assistance Program (SNAP), state general public assistance, or other means-tested benefits.

☐ Yes - Attach documentation of the payments. Skip to Section 3.

☐ No - Continue to Item 3.

3. Are you serving as a Peace Corps volunteer?

☐ Yes - Attach documentation certifying your period of service. Skip to Section 3.

☐ No - Continue to Item 4.

4. Do you work full time? Full-time employment means working 30 or more hours per week in a position expected to last at least 3 consecutive months.

☐ Yes - Continue to Item 5.

☐ No - You are not eligible for this deferment.

5. What is your monthly income? You must attach documentation of your monthly income. Monthly income is either (you choose):

- Your gross taxable income from all sources or
- One-twelfth of the Adjusted Gross Income from your most recent federal income tax return.

6. What is your family size? \_\_\_\_\_

Family size includes:

- You.
- Your spouse.
- Your children if they receive more than half of their support from you, including unborn children who will be born during the deferment period; and
- Other people if, at the time you request this deferment, they live with you, receive more than half their support from you, and will continue to receive this support from you during the deferment period. Support includes money, gifts, loans, housing, food, clothes, car, medical and dental care, and payment of college costs.


7. Is the amount you reported in Item 5 less than 150% of the poverty guideline for your family size and state of residence (see Table 2 in Section 5)?

☐ Yes - Continue to Section 3.

☐ No - You are not eligible for this deferment.

Page 1 of 4

EHD



PUBLIC SERVICE LOAN FORGIVENESS (PSLF) & TEMPORARY EXPANDED PSLF (TEPSLF) CERTIFICATION & APPLICATION

William D. Ford Federal Direct Loan (Direct Loan) Program

OMB No. 1845-0110

Form Approved

Exp. Date 08/31/2023

**WARNING:** Any person who knowingly makes a false statement or misrepresentation on this form or on any accompanying document is subject to penalties that may include fines, imprisonment, or both, under the U.S. Criminal Code and 20 U.S.C. 1097.

**SECTION 1: BORROWER INFORMATION**

Please enter or correct the following information.

☐ Check this box if any of your information has changed.

SSN \_\_\_\_\_

Date of Birth \_\_\_\_\_

Name \_\_\_\_\_

Address \_\_\_\_\_

City \_\_\_\_\_ State \_\_\_\_\_ Zip Code \_\_\_\_\_

Telephone - Primary \_\_\_\_\_

Telephone - Alternate \_\_\_\_\_

Email \_\_\_\_\_

**SECTION 2: BORROWER REQUEST, UNDERSTANDING, AND CERTIFICATION**

For more information on PSLF, visit [StudentAid.gov/pslftips](http://StudentAid.gov/pslftips). To apply online, visit [StudentAid.gov/PSLF](http://StudentAid.gov/PSLF).

**Requirement (1)** that the U.S. Department of Education (the Department) determine whether I qualify for PSLF or TEPSLF, and discharge any qualifying loans that I have, and **(2)** if none of my loans qualify for PSLF or TEPSLF when I submit this form, determine how many qualifying payments I have made towards PSLF and TEPSLF.

☐ I just want to find out how many qualifying payments I have made or if my employer is a qualified employer.

☐ I believe I qualify for forgiveness under PSLF or TEPSLF right now.

☐ If I indicated that I believe I qualify for forgiveness now, I want a forbearance while my application is being processed, but I understand that periods of forbearance do not count towards forgiveness.

**I understand this:**

- To qualify for forgiveness, I must have made 120 qualifying payments on my Direct Loans while employed full-time by a qualifying employer. Neither the 120 qualifying payments nor employment have to be consecutive.
- To qualify for forgiveness, I must be employed full-time by a qualifying employer when I apply for and get forgiveness.
- By submitting this form, my student loans held by the Department may be transferred to MOHELA.
- If the Department determines that I appear to be eligible for forgiveness, the Department may contact my employer before granting forgiveness to ensure that I continue to work for the employer.
- If I am eligible for forgiveness, the amount forgiven will be the principal and interest that was due on my eligible Direct Loans when I made my final qualifying payment. Any amount that I pay on those loans after I have made my final qualifying payment will be treated as an overpayment. I must continue to make payments on any of my other loans.
- If I am not eligible for forgiveness, I will be notified of the determination, why it was made, and how many qualifying payments I have made towards PSLF and TEPSLF.


**I certify** that all of the information I have provided on this form and in any accompanying document is true, complete, and correct to the best of my knowledge and belief and that if I cease to be employed by a qualifying employer after I submit this application, but before forgiveness is granted, I will notify the Department (see Section 7) immediately.

☐ Check this box if you cannot obtain certification from your employer because the organization is closed or because the organization has refused to certify your employment. The Department will follow up to assist you in getting documentation of your employment. **Complete Section 3, but do not complete Section 4.**

Borrower's Signature \_\_\_\_\_ Date \_\_\_\_\_

Page 1 of 6

PSLF



REAFFIRMATION AGREEMENT

William D. Ford Federal Direct Loan (Direct Loan) Program  
Federal Family Education Loan (FFEL) Program

OMB No. 1845-0133

Form Approved

Expiration Date:  
07/31/2021

**WARNING:** Any person who knowingly makes a false statement or misrepresentation on this form or on any accompanying document is subject to penalties that may include fines, imprisonment, or both, under the U.S. Criminal Code and 20 U.S.C. 1097.

**SECTION 1: BORROWER IDENTIFICATION**

Please enter or correct the following information.

☐ Check this box if any of your information has changed.

SSN \_\_\_\_\_

Name \_\_\_\_\_

Address \_\_\_\_\_

City, State, Zip Code \_\_\_\_\_

Telephone - Primary ( ) - \_\_\_\_\_

Telephone - Alternate ( ) - \_\_\_\_\_

Email (Optional) \_\_\_\_\_

**SECTION 2: INFORMATION ABOUT ELIGIBILITY**

You have lost eligibility for federal student financial aid because you inadvertently received a Direct Loan or FFEL program loan that caused you to exceed the annual or aggregate loan limit ("overborrowing"), as shown below. To get complete information about your loan(s) or contact information regarding your loan holder(s), you may visit the National Student Loan Database System (NSLDS) at [NSLDS.ed.gov](http://NSLDS.ed.gov).

To regain eligibility for federal student financial aid, you may either:

- Repay the excess loan amount now, in which case you should contact your loan holder for instructions and not complete this form; or
- Agree to repay the excess according to the terms and conditions of your promissory note ("reaffirmation"), in which case you should review the information in Section 3 and then sign and return this form to your loan holder.

**NOTE:** If the overborrowing was caused by more than one loan and the loans are held by different loan holders, a separate form will need to be submitted to each loan holder.

**SECTION 3: SCHOOL AND LOAN INFORMATION (TO BE COMPLETED BY THE SCHOOL)**

**NOTE:** If NSLDS shows that a student consolidated the loan(s) that caused the inadvertent overborrowing into a Direct or FFEL Consolidation Loan, no further action on the part of the borrower is needed. By signing the consolidation loan promissory note, the borrower agreed to repay any excess loan amount.

Name of Institution \_\_\_\_\_ OPEID \_\_\_\_\_

Address \_\_\_\_\_ City, State, Zip Code \_\_\_\_\_

Name/Title of School Official \_\_\_\_\_ Telephone \_\_\_\_\_

Loan Type	First Disb. Date	Disbursed Amount	Excess Amount	Loan Holder/Servicer

Page 1 of 3

REA

# TRY IT OUT

[illegible]

# TRY IT OUT

<b>Direct Consolidation Loan Application and Promissory Note</b> <b>William D. Ford Federal Direct Loan Program</b>		<small>OFFICIAL USE ONLY</small> <small>FOR DEPARTMENT OF EDUCATION</small>
<b>WARNING:</b> This program is available only to borrowers who are eligible to receive federal student aid. The borrower's borrowing eligibility is subject to review. Failure to follow form instructions, or failure to meet the U.S. Federal Debt and (2011) S.D. 2011		
<b>DEBTOR'S INFORMATION</b>		
Read the instructions for Completing the Direct Consolidation Loan Application and Promissory Note.		
<b>NOTE:</b> PAGE 1 INFORMATION MUST BE SUBMITTED FOR YOUR LOAN APPLICATION TO BE PROCESSABLE		
<b>DEBTOR'S INFORMATION</b>		
1. Last Name	UHWH GQLAYq	first name yH
2. Former Name(s)	Xh(* /ixUN8:   bjqta	
3. Social Security Number(s)	J+INLTSXxc18	
4. Date of Birth	09W17(DIK 6+10U<(	
5. Present Address (Street, City, State, Zip Code) (If P.O. Box, use general address, see form instructions)	P.H~V	
6. Date of Birth (Continuation Number)	PGRq	
7. Email Address (optional)	41~0s1~0911q	
8. Email Address (Work and Home)	QI,QIPyoNQ	
9. Employer's name and address (Street, City, State, Zip Code)	duKeNFRBpHkqOhH&-	
10. Work Area Code/Telephone Number	XIVTw\lpB,G8.)zKv&(	
<b>DEBTOR'S INFORMATION</b>		
List the names and addresses of all persons who are co-borrowers on the loan. Add any additional names and addresses of all persons who are co-borrowers.		
11. Last Name	uv&18/X	first name H6V
12. Date of Birth	T'Z(9	
13. Present Address (Street, City, State, Zip Code)	TB-(M6m	
14. Email Address (optional)	P2<.=	
15. Email Address (Work and Home)	8bW''0	
16. Employer's name and address (Street, City, State, Zip Code)	yC~HVJHvYD	
17. Work Area Code/Telephone Number	3P's&9	
18. Last Name	3HJ8~3+vy*IN18;md~d	first name eR\gn
19. Present Address (Street, City, State, Zip Code)	HJHJWpE(x1)	
20. Email Address (optional)	691w0>>7d#eol\3B9(	
21. Email Address (Work and Home)		
22. Employer's name and address (Street, City, State, Zip Code)		
23. Work Area Code/Telephone Number		
Signature of Debtor		

# DLC



# TRY IT OUT

[illegible]

# TRY IT OUT

[illegible]

PSLF

# DOCUMENT CLASSIFICATION

## WHY IT MATTERS?

July 2022

- Received 421,000 Documents

Volume is seasonal

- Overtime/Temp Workers

Monthly Volume Pre CARES Act

- >1,000,000

## THE OPPORTUNITY

Scalability

- Expand/Shrink to Volume

Cost of Servers/Compute

- < Hourly wage

Human Error

- Human Error is less predictable

# DATA

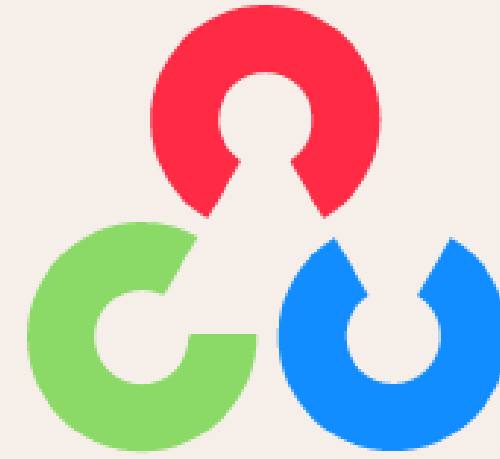
Dynamically Generated Dataset

200x200 Pixel Images

204,000 Images

- 36,000/Class(6)

# T O O L S



OpenCV

- Image Processing



Albumentations

- Image Augmentation



NumPy

- Data Handling



Sklearn

- Analysis/Modeling



# MODELING

## Bernoulli Naive Bayes

- Designed for Sparse Binary Data
- Assigns Probability to for each pixel
- Fast Training/Inference
- Online Training

## Our Approach

- 6 Models (1 Per Class)
- Models "Vote" On Most Likely Class
- Contested Votes Flagged as UNK

## Performance

- Precision: 0.53
- Recall: 0.5
- F1: 0.5

		Predicted					
Actual		686	83	56	23	81	71
		213	534	40	40	214	49
		105	5	470	81	305	34
		105	52	162	402	249	30
		109	42	150	62	590	47
		217	145	96	54	160	328
		DLC	TLF	IDR	EHD	PSLF	REA

## Production vs Training.

- # How can our model handle 1/4 unseen documents?

Page 2 of 4Page 3 of 3

- Precision: 0.86
- Recall: 0.87
- F1: 0.86



# DEMO



Created with Gradio  
Hosted with HuggingFace



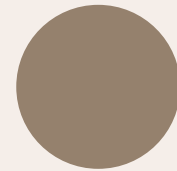
<https://huggingface.co/spaces/PatrickTyBrown/LoanDocumentClassifier>

# PROOF OF CONCEPT

Effective.



Simple.



Scalable.



# NEXT STEPS

## Baseline

- Current Model will function as a baseline

## Neural Networks

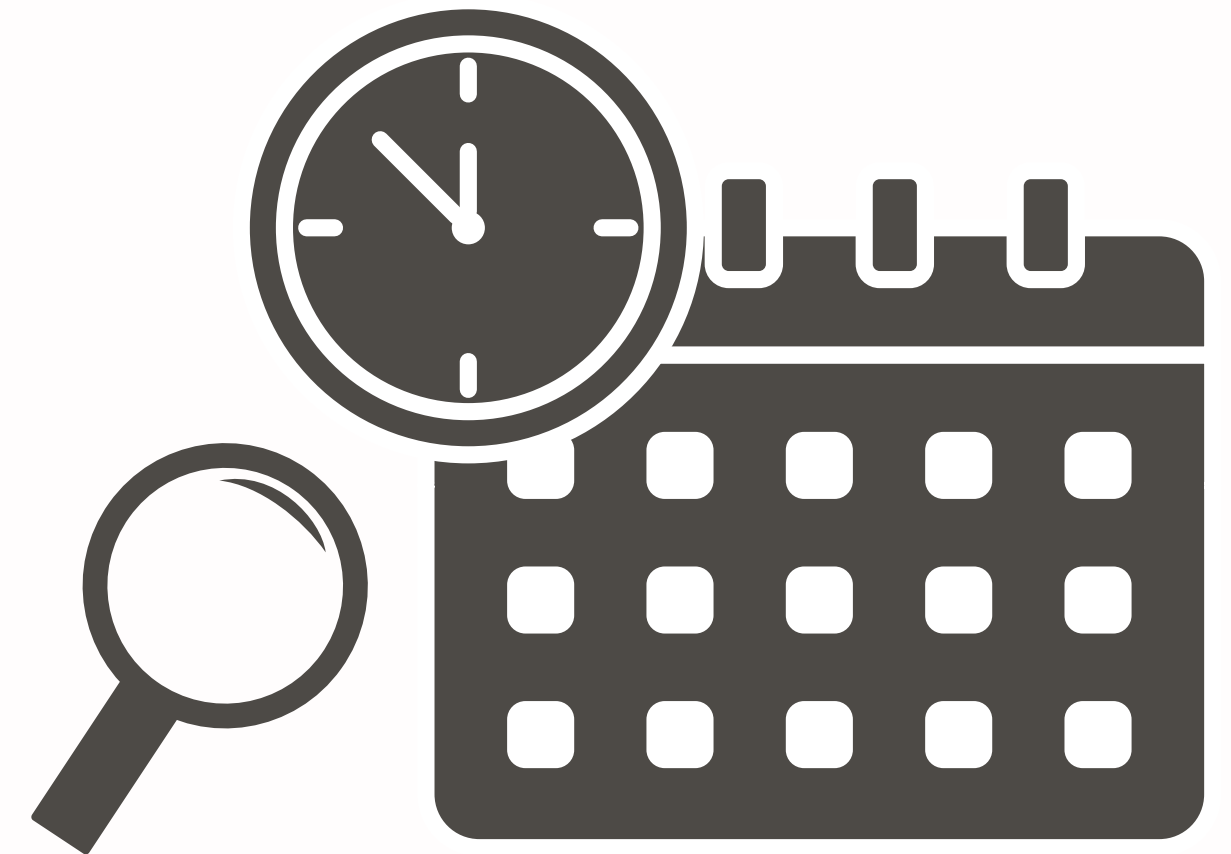
- Explore more complex modeling

## Preprocessing

- Utilize more advanced CV techniques for preparing documents

## Multimodality

- Utilize OCR for capturing text
- Incorporate text data into modeling classification





# APPENDIX

Feel free to ask me any questions  
on inspiration or implementation!

## Contact

Email: [patrick.ty.brown@gmail.com](mailto:patrick.ty.brown@gmail.com)

Linkedin: <https://www.linkedin.com/in/patrick-ty-brown/>

# C O N T E N T S

-Data Generation

-Data Cleaning

-Unsuccess stories  
(Attempted Strategies)

-Model Selection

-Model Interpretation

# GENERATING THE DATASET

Dynamically Generated Dataset  
204,000 (36,000/Class)

Step 1

- Take Base Form Image

Step 2

- Add Random Text

Step 3

- Place Image on Background

Step 4

- Random Augment Image

Base Form

Direct Consolidation Loan Application and Promissory Note  
William D. Ford Federal Direct Loan Program

WARNING: Any person who knowingly makes a false statement or misrepresentation on this form or any accompanying document is subject to penalties that may include fines, imprisonment, or both, under the U.S. Criminal Code and 20 U.S.C. 1097.

BEFORE YOU BEGIN

Read the Instructions for Completing the Direct Consolidation Loan Application and Promissory Note.  
NOTE: PAGES 1 THROUGH 5 MUST BE SUBMITTED FOR YOUR LOAN APPLICATION TO BE PROCESSED.

**BORROWER INFORMATION**

1. Last Name: \_\_\_\_\_ First Name: \_\_\_\_\_ Middle Initial: \_\_\_\_\_  
2. Former Name(s): \_\_\_\_\_  
3. Social Security Number: \_\_\_\_\_  
4. Date of Birth: \_\_\_\_\_  
5. Permanent Address (Street, City, State, Zip Code) (If P.O. box or general delivery, see instructions): \_\_\_\_\_  
6. Area Code/Telephone Number: \_\_\_\_\_  
7. Email Address (optional): \_\_\_\_\_  
8. Driver's License State and Number: \_\_\_\_\_  
9. Employer's Name and Address (Street, City, State, Zip Code): \_\_\_\_\_  
10. Work Area Code/Telephone Number: \_\_\_\_\_

**REFERENCE INFORMATION**

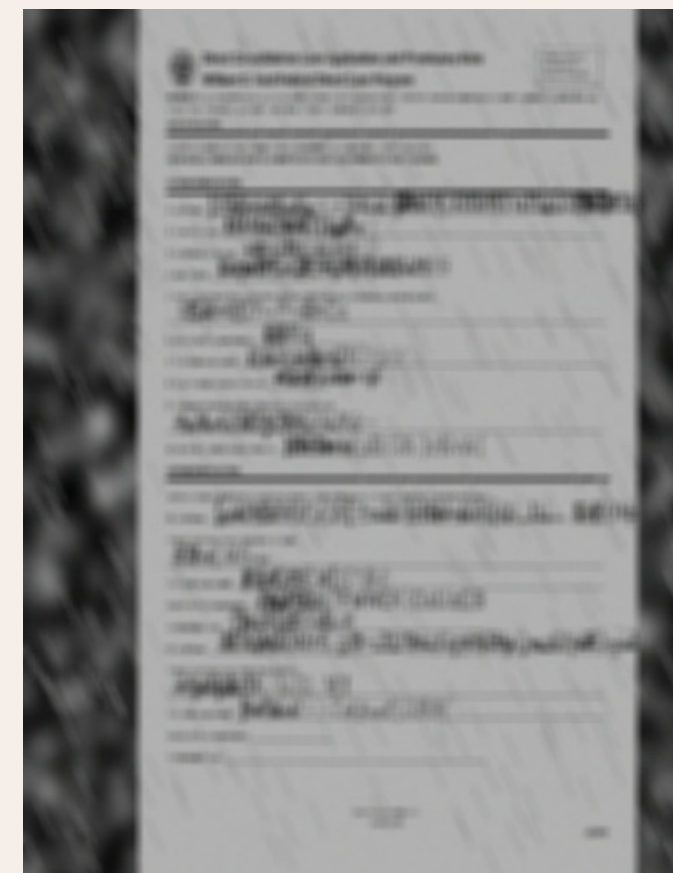
List two persons with different U.S. addresses who do not live with you and who have known you for at least three years.

11. Last Name: \_\_\_\_\_ First Name: \_\_\_\_\_ Middle Initial: \_\_\_\_\_  
Permanent Address (Street, City, State, Zip Code): \_\_\_\_\_  
Email Address (optional): \_\_\_\_\_  
Area Code/Telephone Number: \_\_\_\_\_  
Relationship to You: \_\_\_\_\_  
12. Last Name: \_\_\_\_\_ First Name: \_\_\_\_\_ Middle Initial: \_\_\_\_\_  
Permanent Address (Street, City, State, Zip Code): \_\_\_\_\_  
Email Address (optional): \_\_\_\_\_  
Area Code/Telephone Number: \_\_\_\_\_  
Relationship to You: \_\_\_\_\_

SUMMARY PAGES 1 THROUGH 5  
Page 1 of 16

12/2019

Processed



# IMAGE PROCESSING

850 X 1100 RGB Image

- 2,805,000 Data Points

Step 1

- Convert to Image Grayscale

Step 2

- Apply Thresholding

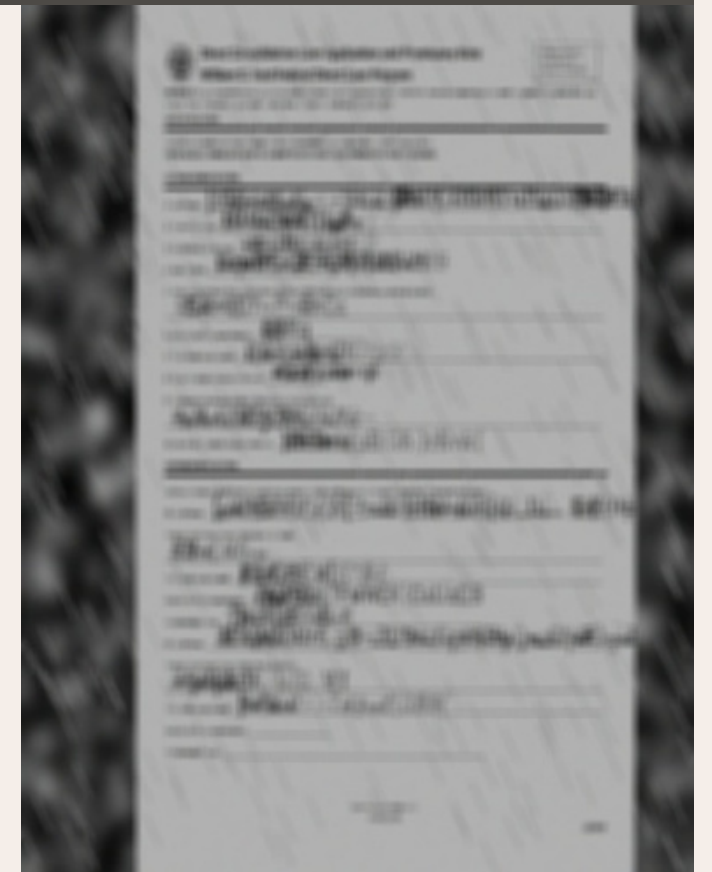
Step 3

- Resize Image to 200x200

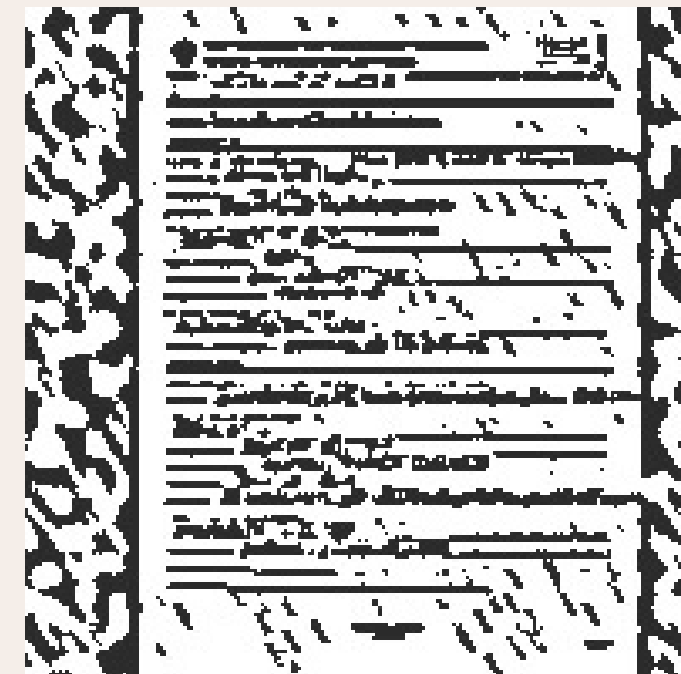
Step 4

- Flatten to 40,000 Numpy Array

Augmented  
Image



Processed



Data

[1,0,0,...,1,0,1]

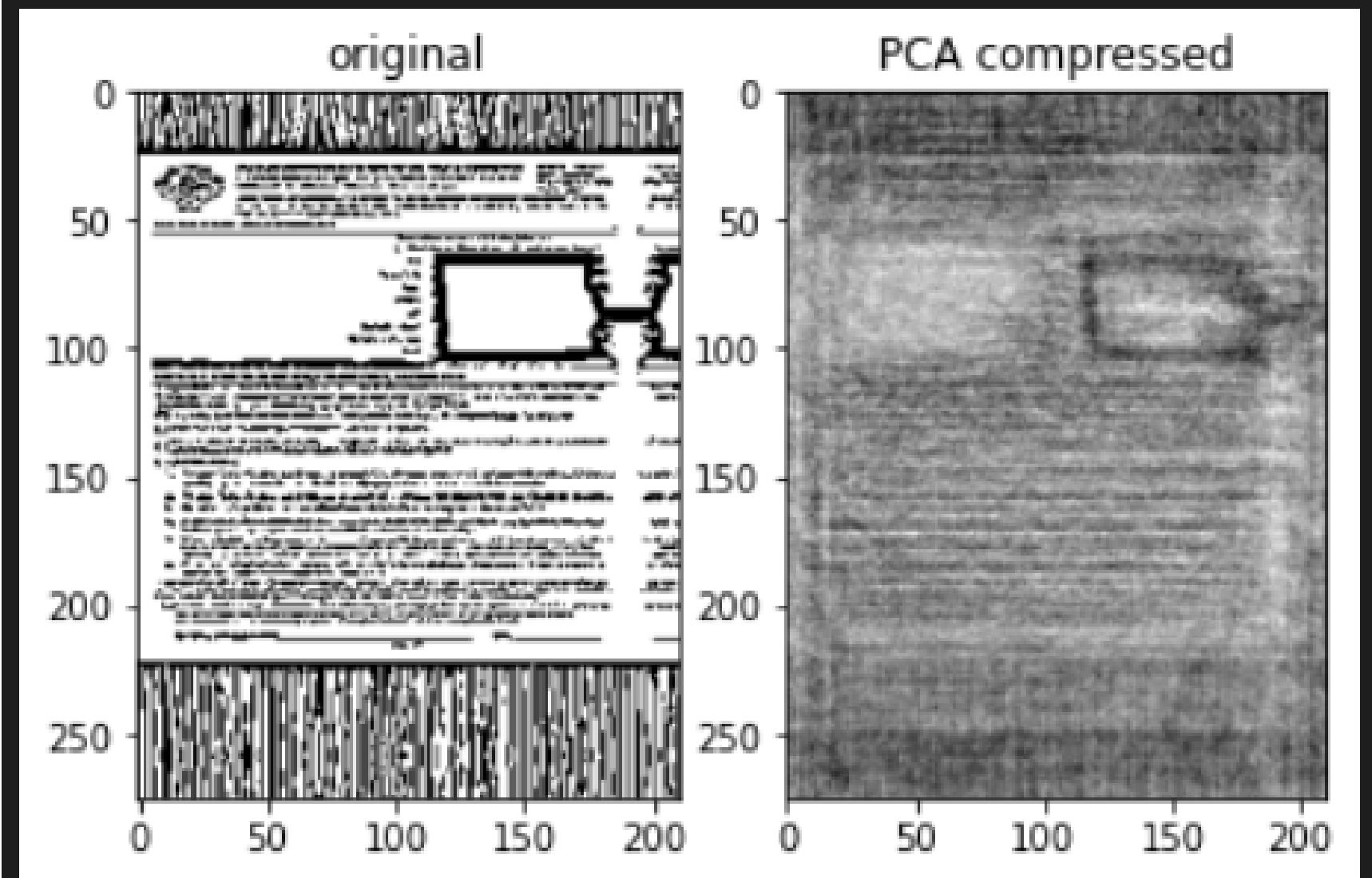
# THE UNSUCCESS STORY (THE ORIGINAL ATTEMPT)

1

## Principle Components Analysis (PCA)

- Used to reduce dimensions of data to  $<500$  Features
- 360 Features most effective
  - 88% Accuracy

reduced shape: (1, 360)  
recovered shape: (1, 57750)



# THE UNSUCCESS STORY

## (THE ORIGINAL ATTEMPT)

2

# Logistic Regression

- Naive model for baseline classification metrics
- Predicting 5 Classes
- Performance with tuning
  - F1: 0.225

Actual

Predicted				
2162	2033	1902	1918	1985
1840	2305	1941	1966	1948
1940	2027	2243	1851	1939
1917	1954	1948	2247	1934
1917	1954	1948	2247	1934
DLC	TLF	IDR	EHD	PSLF



# THE UNSUCCESS STORY

## (THE ORIGINAL ATTEMPT)

3

K Nearest Neighbors

- F1: 0.215

Decision Tree

- F1: 0.204

Random Forest

- F1: 0.223

Gaussian Naive Bayes

- F1: 0.228

Support Vector Machine

- F1: 0.224

XGBoost

- F1: 0.228

# MODEL SELECTION

## Multiclass (MC) Model vs One Vs Rest (OVR)

- MC models performed 10-15% better

## Why OVR then?

- OVR Ensemble introduced opportunity to flag contested documents
  - After filtering flagged results OVR performed slightly better
- OVR can recognize negative class

## Exploring Various Image Sizes

42 x 65

- F1: 0.708

85 x 110

- F1: 0.68

200 x 200

- F1: 0.774

280 x 360

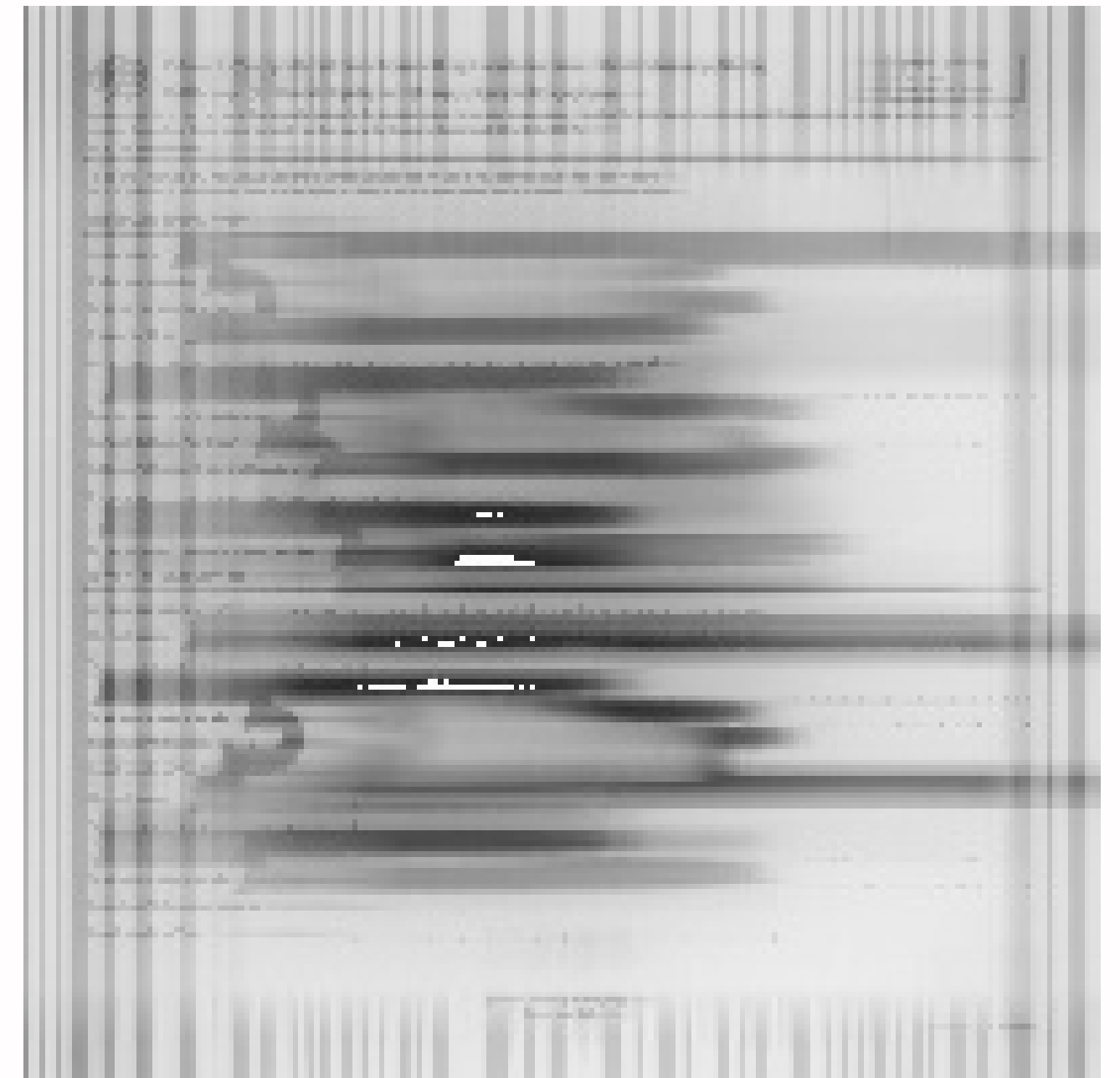
- F1: 0.773

# MODEL INTERPRETATION

1

Where is the model looking?

- Reconstructed Images out of model created probabilities
- Models roughly learned the template forms through the training process



# MODEL INTERPRETATION

2

