Loans @ a Click

Digital Access | **Data Driven** | **Enabled by Technology**

Why does FinTech exist?

She is the victim of a Jaan Pehchaan model ...

... and we are solving her problem

- Runs a small business
- Based in Udaipur
- No CIBIL record, but occasionally needs financing
- Is not in the ‘privilege segment’ of any bank
- ‘Phones a friend’ when she needs money because it is quick (not cheap)

And there are many like him…

…48 mn to be precise

- 94% do not have a credit score
- >50% need but do not have access to formal credit

PROBLEM

ACCESS

+

ASSESSMENT

+

EFFICIENCY

SOLUTION

PARTNER NETWORK

+

ALTERNATIVE DATA BASED SCORING

+

DIGITAL PROCESSING

Digital Access	Data Driven	Enabled by Technology

Realizing Our Dream driven by Data-driven decision making

PRESENT

Manual intervention is needed for credit assessment

TAT~48-72 hrs

1. Loan Application

3. Loan Disbursed

2. Manual Credit Assessment

FUTURE

Automated Credit Assessment enabling *"Loans at a click"*

TAT~1 hr

1. Loan Application

3. Loan Disbursed

2. Automated Credit Assessment

FINTECH AND FLEXILOANS

WHY SHOULD YOU BE INTERESTED

How to Improve Customer Experience through faster processing?

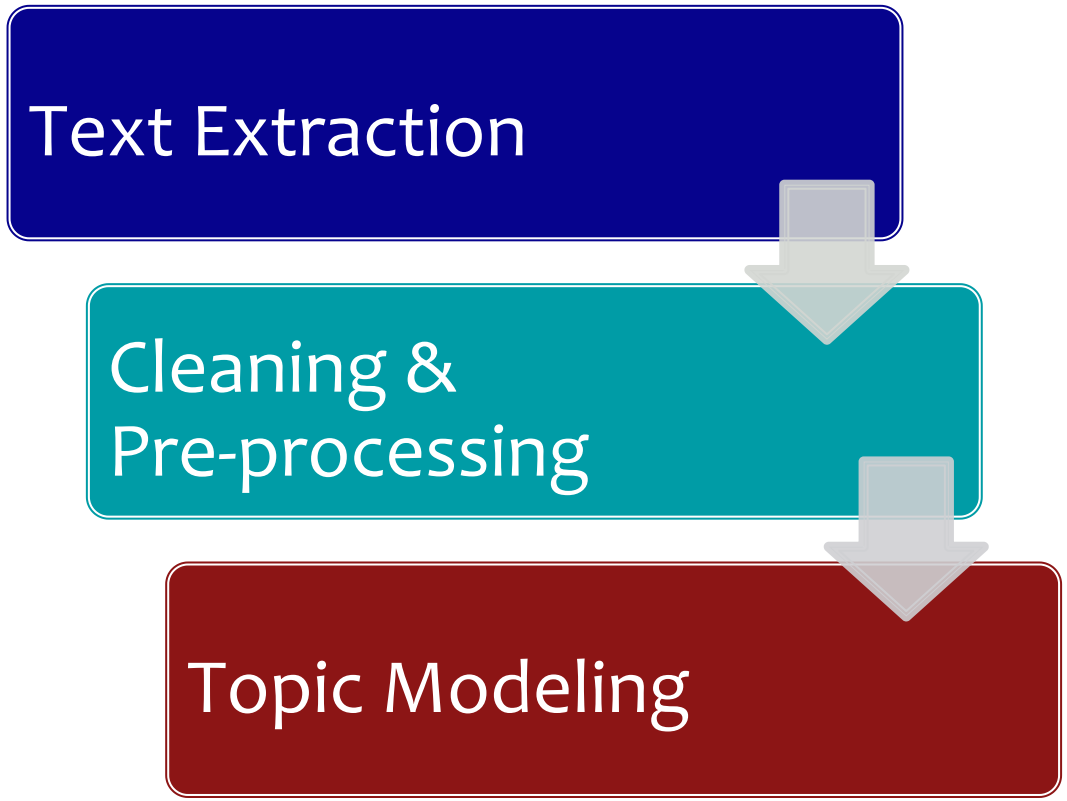
The Business Problem

- To improve a customer's application processing time by analyzing the remarks entered during Credit Appraisal process and finding appropriate ways to improve the process

Proposed Solution

- Extract the text data given at every stage of application processing
- Cleaning and Pre-processing the text data using NLP
- Finding the main cause of hindrance in the loan process by using topic modeling

Solution Overview



Proposed solution framework

Phase I : Text Extraction

TOKENIZATION : It breaks unstructured data, text, into chunks of information which can be counted as discrete elements

```
[update, lead, name, customer,
                [done, analysis, ba
  [certificate_of_registration,
[answer, sms, mail, dropped., h

 pd_comments, dtype: object
```

Phase II : Data Cleaning & Preprocessing

STOP WORD REMOVAL : Certain parts of English speech, like conjunctions (“for”, “or”) or the word “the” are meaningless to a topic model.These terms are called stop words and need to be removed from our token list.

STEMMING : Used to reduce topically similar words to their root. Porter stemmer returns the string parameter in stemmed form.

Phase III : Topic Modeling

WORD CLOUD : It will show the most occuring topic or word in the given text document.

LDA FOR TOPIC MODELING : Latent Dirichlet allocation (LDA) is particularly useful for finding reasonably accurate mixtures of topics within a given document set.

```
ldamodel = gensim.models.ldamodel.LdaModel(corpus, num_topics=10,
```

```
print(ldamodel.print_topics(num_topics=8, num_words=4))
```

```
[(9, u'0.036*"bank_statement" + 0.020*"account" + 0.016*"data" +
0.021*"bank_statements" + 0.021*"financials"'), (0, u'0.036*"bank
0.014*"card"'), (3, u'0.037*"business" + 0.035*"kyc" + 0.028*"per
atement" + 0.032*"aug" + 0.025*"17" + 0.016*"account"'), (4, u'0.
ata" + 0.014*"ask"'), (2, u'0.049*"bank_statement" + 0.027*"name"
011*"company" + 0.011*"data" + 0.009*"applicant" + 0.009*"copy"')
```

Implementing the Solution Framework

Text Remarks

```
: X.head()

: 0    Please update  the lead by name of the custome...
  1    we have done the analysis for his Bank Bank , ...
  2    Require certificate_of_registration and Bank C...
  3    no answer, sms and mail dropped.&&Aniket Date_...
  4    pan_card not updated on LOS&&Priyanka Gaikwad ...
  Name: pd_comments, dtype: object
```

Pre-processing Text

```
X = preprocess(X)
X

/home/admin1/.local/lib/python2.7/site-packages/ipykernel
ailed to convert both arguments to Unicode - interpreting

0       [update, lead, name, customer, business, name,...
1                  [done, analysis, bank_statement, aug]
2      [certificate_of_registration, current, account]
3      [answer, sms, mail, dropped., hdfc, aug, state...
4                                  [pan_card, updated]
5                  [confirm, flagship, model, product]
6      [pan_card, company, kyc, docs, kyc, shareholde...
```

Implementing LDA Model

```
ldamodel = gensim.models.ldamodel.LdaModel(corpus, num_topics=10, id2word = dictionary, passes=100)
```

```
print(ldamodel.print_topics(num_topics=8, num_words=4))
```

```
[(9, u'0.036*"bank_statement" + 0.020*"account" + 0.016*"data" + 0.015*"docs"'), (7, u'0.029*"fy" + 0.022*"itr" +
0.021*"bank_statements" + 0.021*"financials"'), (0, u'0.036*"bank_statement" + 0.022*"address" + 0.014*"pattern" +
0.014*"card"'), (3, u'0.037*"business" + 0.035*"kyc" + 0.028*"per" + 0.024*"bank_statement"'), (8, u'0.067*"bank_st
atement" + 0.032*"aug" + 0.025*"17" + 0.016*"account"'), (4, u'0.044*"bank_statement" + 0.024*"customer" + 0.019*"d
ata" + 0.014*"ask"'), (2, u'0.049*"bank_statement" + 0.027*"name" + 0.019*"pdf_format" + 0.019*"proper"'), (6, u'0.
```

Word Cloud

```
wordCloud(X)
```

THANK YOU !!

For Any queries please reach out to me at :

Email: nitika.goel@flexiloans.com