Problem Statement

Many of us have dreams what if I become the founder of a start-up and can lead it to become a big business like Microsoft, or Facebook/Amazon. But again there are many people who cheat people to run their business, probably everyone has heard about cheat-funds. This dataset contains a lot of data and the 'LICENSE STATUS' column which depicts whether the person has been able to get a business license or not. Based on the given features you have to predict whether a person will be given a license or not to run his business. The 'LICENSE STATUS' contains the following categories:

1. AAC (license was cancelled during term)

2. AAI (license was issued)

3. INQ (license requires inquiry)

4. REA (license revocation has been appealed)

5. REV (license was revoked)

You have to predict the column "LICENSE STATUS" and please submit in the format given in the "sample_submissions.csv" file.

Evaluation Criterion

The evaluation criteria will be the f1- score with 'weighted' average. You can read more about the metrics here.

$$F1\ score = \frac{Precision * Recall}{Precision + Recall}$$

The weighted precision or recall is defined as:

$$\frac{1}{\sum_{l \in L} |\hat{y}_l|} \sum_{l \in L} |\hat{y}_l| \phi(y_l, \hat{y}_l)$$

- L is the set of labels
- $\hat{y}$ is the true label
- y is the predicted label
- $\hat{y}_l$ is all the true labels that have the label l
- $|\hat{y}_l|$ is the number of true labels that have the label l
- $\phi(y_l, \hat{y}_l)$ computes the precision or recall for the true and predicted labels that have the label l. To compute `precision`, let $\phi(A, B) = \frac{|A \cap B|}{|A|}$. To compute `recall`, let $\phi(A, B) = \frac{|A \cap B|}{|B|}$.

Download dataset

Sl. No.	Column Label	Column Description
1	ID	A calculated unique ID for each record.
2	LICENSE ID	An internal database ID for each record. Each license can have multiple records as it goes through renewals and other transactions. See the LICENSE NUMBER field for the number generally known to the public and used in most other data sources that refer to the license.
3	ACCOUNT NUMBER	The account number of the business owner, which will stay consistent across that owner's licenses and can be used to find the owner in the Business Owners dataset.
4	SITE NUMBER	An internal database ID indicating the location of this licensed business to account for business owners with more than one location.
5	LEGAL NAME	Legal document-wise name
6	DOING BUSINESS AS NAME	Name he has given in his company
7	ADDRESS	Address of the business office

8	CITY	City of the business office
9	STATE	State of the business office
10	ZIP CODE	Zip code of the business office
11	WARD	Ward of the business office
12	PRECINCT	The precinct within the ward where the business is located. Note the same precinct numbers exist in multiple wards.
13	WARD PRECINCT	The ward and precinct where the business is located. This column can be used to filter by precinct more easily across multiple wards.
14	POLICE DISTRICT	The police district where the business is located
15	LICENSE CODE	The code for the type of license
16	LICENSE DESCRIPTION	Purpose for which license is taken

17	LICENSE NUMBER	The license number known to the public and generally used in other data sources that refer to the license. This is the field most users will want for most purposes. Each license has a single license number that stays consistent throughout the lifetime of the license. By contrast, the LICENSE ID field is an internal database ID and not generally useful to external users.
18	APPLICATION TYPE	Type of Application filed
19	APPLICATION CREATED DATE	The date the business license application was created. RENEW type records do not have an application.
20	APPLICATION REQUIREMENTS COMPLETE	For all application types except RENEW, this is the date all required application documents were received. For RENEW type records, this is the date the record was created.
21	PAYMENT DATE	The day payment was completed for license
22	CONDITION APPROVAL	This pertains to applications that contain liquor licenses. Customers may request a conditional approval prior to building out the space.
23	LICENSE TERM START DATE	This is the date from which the license becomes valid.

24	LICENSE TERM EXPIRATION DATE	This is the date when the license is expired and is no longer valid.
25	LICENSE APPROVED FOR ISSUANCE	This is the date the license was ready for issuance. Licenses may not be issued if the customer owes debt to the City.
26	DATE ISSUED	The date license was issued
27	LICENSE STATUS CHANGE DATE	The date when license status was changed if at all it was changed
28	SSA	Special Service Areas are local tax districts that fund expanded services and programs, to foster commercial and economic development, through a localized property tax. In other cities these areas are sometimes called Business Improvement Districts (BIDs). This portal contains a map of all Chicago SSAs
29	LATITUDE	Latitude Location
30	LONGITUDE	Longitude Location
31	LOCATION	Locality of the business

32	LICENSE STATUS	What happened to the license: 1. AAC (license was cancelled during term) 2. AAI (license was issued) 3. INQ (license requires inquiry) 4. REA (license revocation has been appealed) 5. REV (license was revoked)

Sample Submission

You have to predict the column "LICENSE STATUS" and please submit in the format given in the "sample_submissions.csv" file.

```
ID,LICENSE STATUS
1985799-20091116,AAC
1932144-20090416,AAI
71173-20100816,REV
```