Diabot

by
Dave - Julia - Miguel

Motivation & Summary

- Diabetes is the biggest epidemic Australia has ever faced (incl COVID)
- Currently one person is diagnosed with Diabetes every 5 mins
- Type 2 Diabetes affects almost 87% of all people with Diabetes
- Type 2 Diabetes costs Australia over $14.6 billion annually
- And here's the head slapper…

It's preventable!

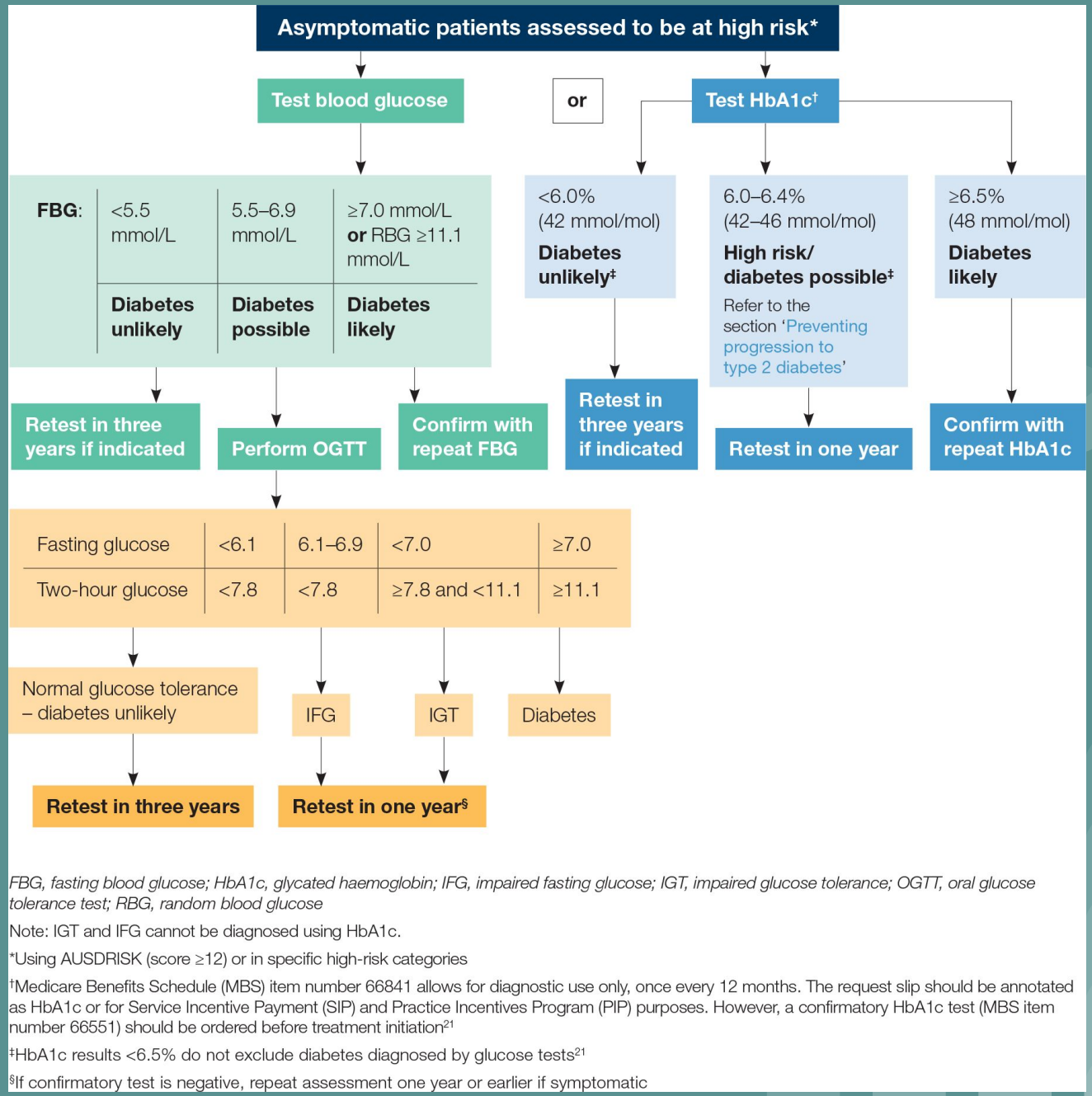
Risk factors for type 2 diabetes include:

- having a family history of type 2 diabetes
- having pre-diabetes
- being above the healthy weight range
- having an inactive lifestyle
- increasing age
- having an Aboriginal or Torres Strait Islander background
- being from a Melanesian, Polynesian, Chinese, Southeast Asian, Middle Eastern or Indian background
- having prior gestational diabetes
- having polycystic ovary syndrome
- taking some types of antipsychotic or steroid medications

RACGP Clinical Pathway for suspected risk of NIDDM

GP: “Alexa the previous Resting Blood Glucose was 13.6 and now the Fasting Blood Glucose is 8.2”

Alexa: “Your patient has Type 2 Diabetes”



FBG, fasting blood glucose; HbA1c, glycated haemoglobin; IFG, impaired fasting glucose; IGT, impaired glucose tolerance; OGTT, oral glucose tolerance test; RBG, random blood glucose

Note: IGT and IFG cannot be diagnosed using HbA1c.

*Using AUSDRISK (score ≥12) or in specific high-risk categories

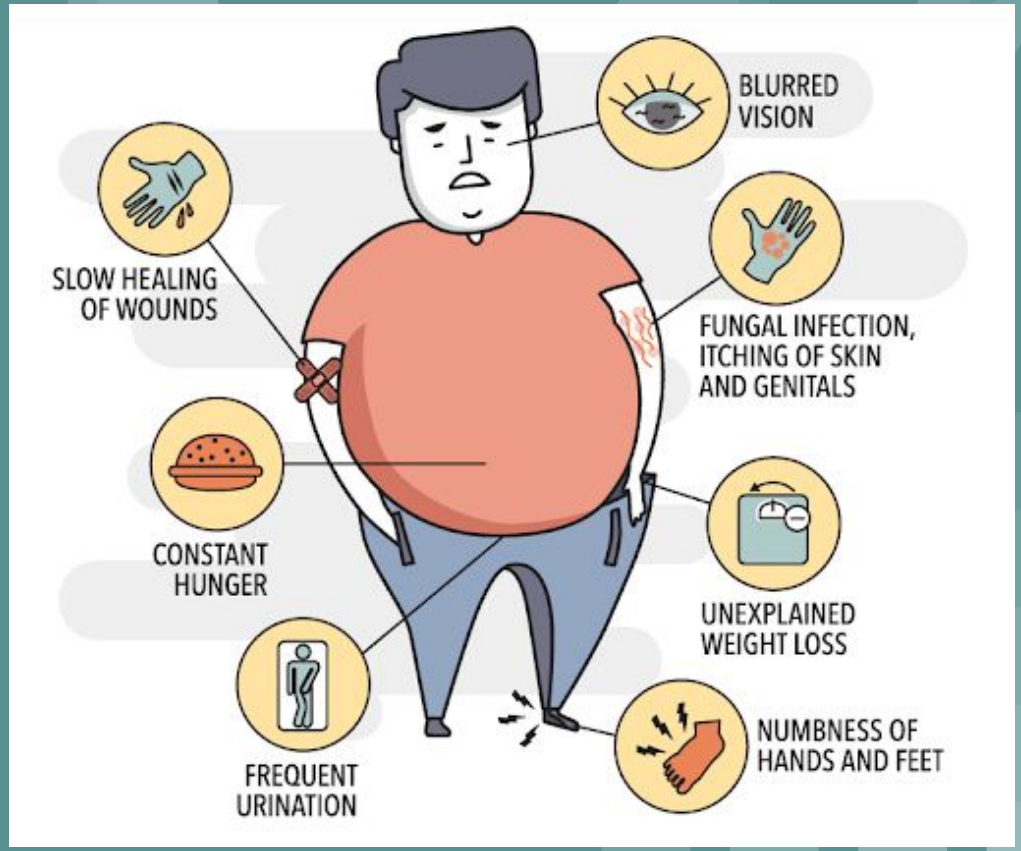
†Medicare Benefits Schedule (MBS) item number 66841 allows for diagnostic use only, once every 12 months. The request slip should be annotated as HbA1c or for Service Incentive Payment (SIP) and Practice Incentives Program (PIP) purposes. However, a confirmatory HbA1c test (MBS item number 66551) should be ordered before treatment initiation[21]

‡HbA1c results <6.5% do not exclude diabetes diagnosed by glucose tests[21]

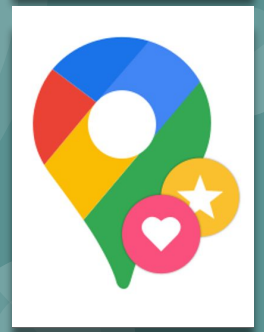
§If confirmatory test is negative, repeat assessment one year or earlier if symptomatic

Symptoms of type 2 diabetes include:

- being thirsty and drinking more than usual
- going to the toilet (to pass urine) more often
- feeling tired and low on energy
- sores or cuts that won’t heal
- blurred vision
- itching and skin infections
- pain or tingling in the legs or feet



Motivation & Summary



- Patients can casually ask our AWS Lex Bot, Diabot which predict the risk of acquiring Type 2 Diabetes and then recommend:
  - Clinical services based on location (pandas) and google reviews (NLP)
- Practitioners can integrate Diabot into their practice management systems and utilise patient health data to support their clinical diagnostic pathways and offer advice on referral services
- Patients or Practitioners can complete a basic questionnaire using JotForms
  - For compatibility with legacy systems
  - As an access consideration (scalable to paper)
- We will use a Machine Learning Algo to predict if the person is at either HIGH or LOW RISK of developing Diabetes and use the Google Maps API to offer recommendations for clinical services based on proximity and sentiment analysis via NLP.

ML Demo

Predicting Diabetes

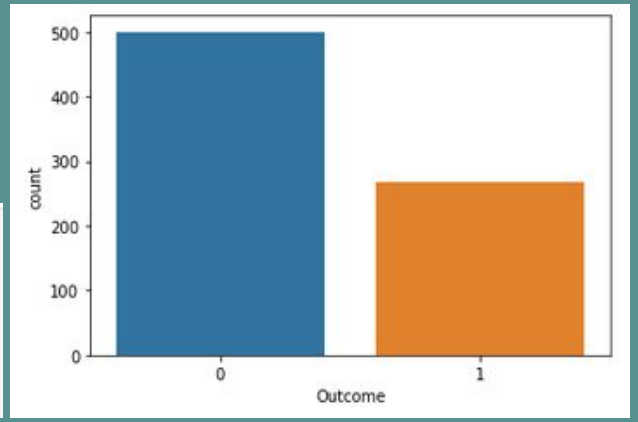
```
diabetes.head()
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

```
diabetes.groupby('Outcome').size()
```

```
Outcome
0    500
1    268
dtype: int64
```

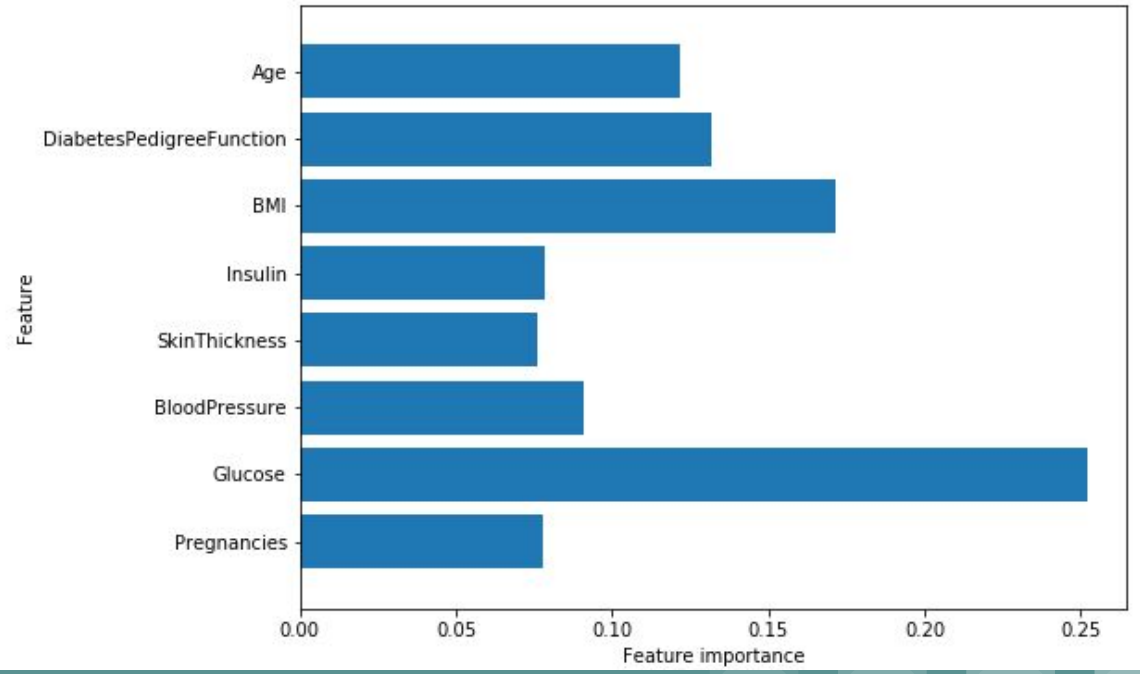
```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(
    diabetes.loc[:,diabetes.columns != 'Outcome'],
    diabetes['Outcome'],
    stratify=diabetes['Outcome'],
    random_state=66
    )
```



The Random Forest gives a lot of importance to the "Glucose" and "BMI" features. The randomness in building the random forest forces the algorithm to consider many possible explanations, the result being that the random forest captures a much broader picture of the data than a single tree and is therefore the best model for our requirements in this project.

Feature importance in Random Forest

```
plot_feature_importances_diabetes(rf)
```



Trialling various models to find the best accuracy on the training and test sets…

- Logistic Regression = 78% / 77%
- Decision Tree = 84% / 75%
- Random Forest = 100% / 76%

Robo Advisor Demo

AWS Lex Bot

First Service - Public

The first service produces a **Risk Score** based on the **Life-Style habits** of the person.

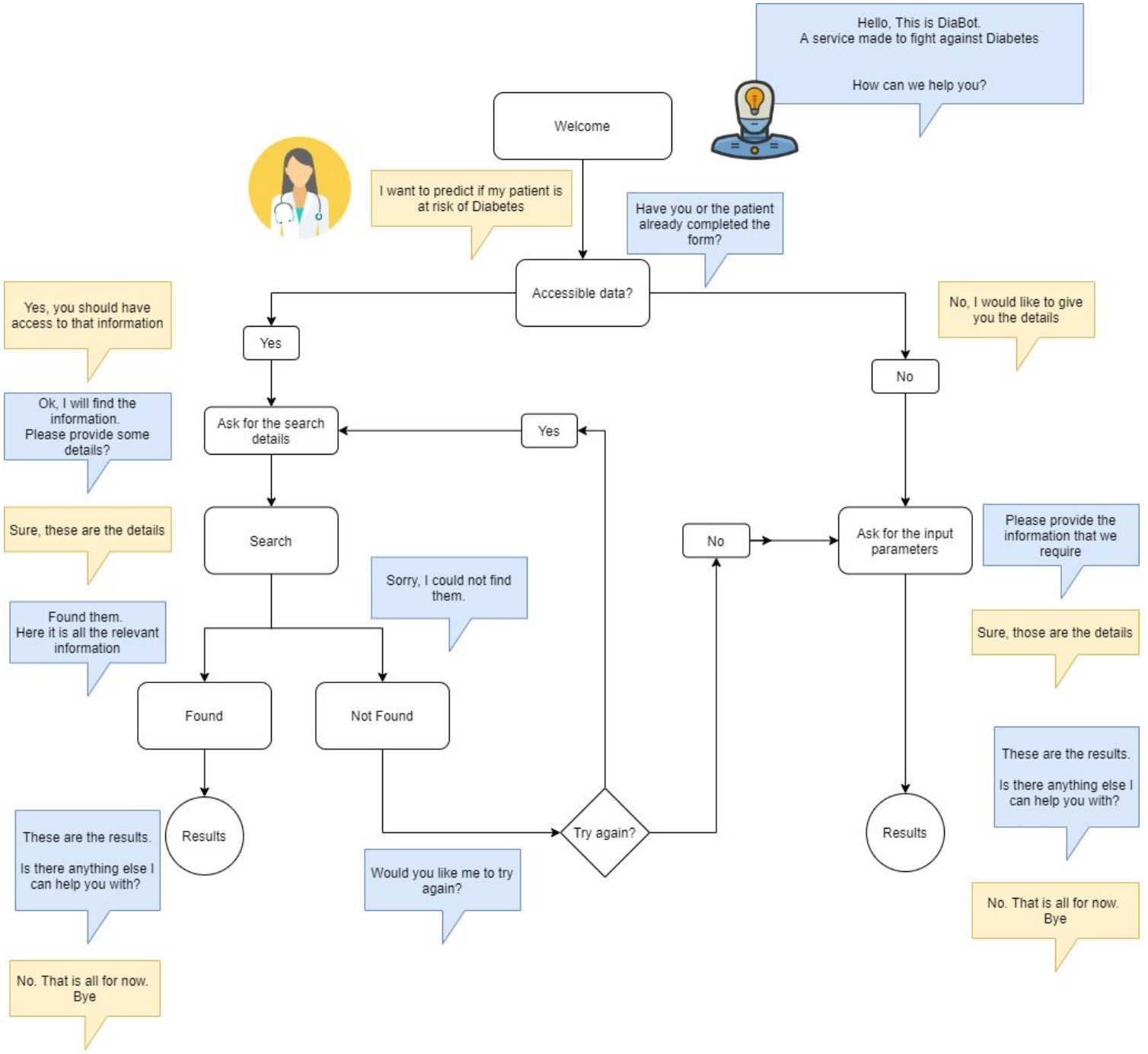
The data ...

Second Service - Practitioners

- **Main features**
  - Designed to leverage ML Models to predict diseases
  - To be released to a more specific audience
  - It is able to communicate with the outside world

The first service produces a **Risk Score** based on the **ML Trained Model** habits of the person.

The data …


Hello, This is DiaBot.
A service made to fight against Diabetes
How can we help you?
Welcome
I want to predict if my patient is at risk of Diabetes
Have you or the patient already completed the form?
Accessible data?
Yes, you should have access to that information
Yes
No, I would like to give you the details
No
Ok, I will find the information. Please provide some details?
Ask for the search details
Yes
Sure, these are the details
Search
No
Ask for the input parameters
Please provide the information that we require
Sorry, I could not find them.
Found them. Here it is all the relevant information
Sure, those are the details
Found
Not Found
These are the results. Is there anything else I can help you with?
Results
Try again?
Results
These are the results. Is there anything else I can help you with?
Would you like me to try again?
No. That is all for now. Bye
No. That is all for now. Bye

Postmortem

Main issues:

- Lex Bot Version. We need to use Lex V2
- Installing packages required for the Lambda Functions
- Security issues will be our biggest hurdle with development of our health bot. More about the HL7 patient data protocol here: https://www.hl7.org/implement/standards/

To do later:

Deployment. Make the bot available in tools such as Slack, Facebook or WhatsApp

Questions