MACHINE LEARNING: WHY SHOULD YOU JUMP ON THE BANDWAGON?

LET'S CONSIDER THE PROBLEM OF

SPAM DETECTION

YOU WORK AT A LARGE EMAIL SERVICE (SAY GMAIL AT GOOGLE)

YOU NEED TO FIGURE OUT A WAY TO TEST IF EMAILS COMING INTO INBOXES ARE **SPAM** OR **HAM** (AS NON-SPAM EMAILS ARE CALLED)

ONE WAY OF DOING THIS -

DEFINE A SET OF **RULES**

"ANY EMAIL FROM A CERTAIN IP ADDRESS OR EMAIL ID IS SPAM"

"ANY EMAIL ID FROM A CONTACT OF A CONTACT IS NOT SPAM"

"ANY EMAIL CONTAINING A CERTAIN SET OF WORDS IS SPAM"

THE PROBLEM WITH A RULE-BASED APPROACH TO SUCH A PROBLEM IS THAT THE RULES ARE RATHER **STATIC** AND CHANGE SLOWLY

WHILE THE BEHAVIOR PATTERNS OF SPAMMERS ARE **DYNAMIC** AND CHANGE SUPER-FAST IN RESPONSE TO THOSE RULES

AN ALTERNATIVE TO A RULE-BASED APPROACH MIGHT BE -

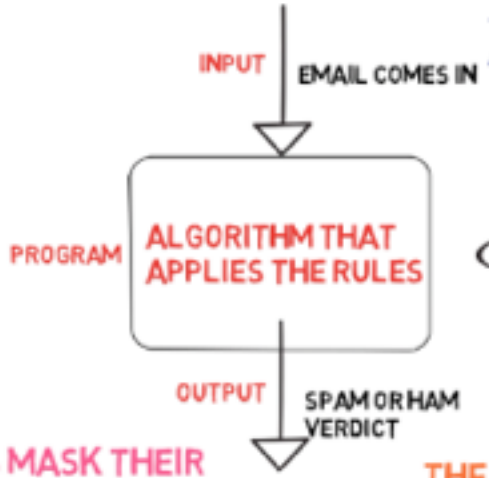
A MACHINE-LEARNING BASED APPROACH

FIGURE OUT PATTERNS IN THE KINDS OF EMAILS THAT ARE EXPLICITLY MARKED AS SPAM BY THE USER

THEN WHEN A NEW EMAIL COMES IN - CHECK TO SEE IF THIS EMAIL CONFORMS TO THOSE SAME PATTERNS

IF YES - MARK IT AS SPAM

THE RULE-BASED APPROACH



THE RULE-BASED APPROACH IS STATIC -

THE ALGORITHM DOES NOT CHANGE BASED ON WHETHER ITS WORKING OR NOT

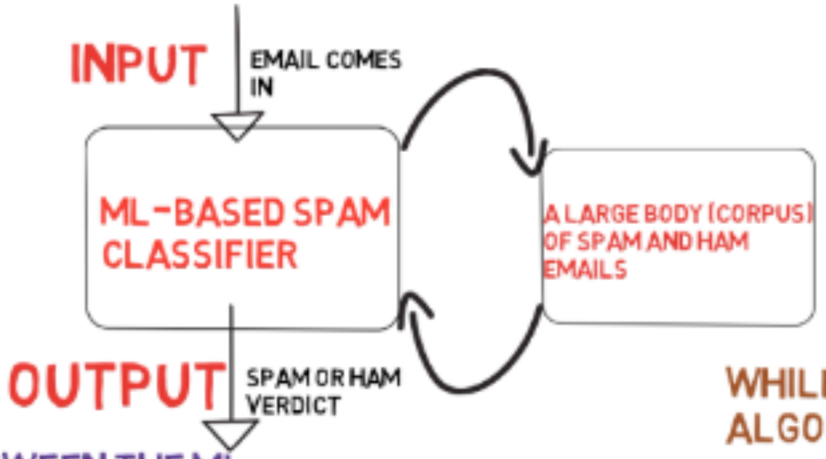
HOW?
REMEMBER THAT USERS GIVE FEEDBACK ON THE QUALITY OF THE ALGORITHM BY MARKING EMAILS AS SPAM, OR SPAM AS "NOT SPAM"

SO, AS SPAMMERS MASK THEIR IP ADDRESSES, CHANGE EMAIL IDS, AND ALTER WHAT THEY ARE SEEKING TO SELL, A RULE-BASED APPROACH WILL SLOWLY BUT INEVITABLY FALL BEHIND

THE STATIC ALGORITHMIC APPROACH IS MISSING OUT ON THE OPPORTUNITY TO IMPROVE ITSELF BASED ON THE FEEDBACK THAT USER ACTIONS PROVIDE

BEHIND

THE ML-BASED APPROACH



INPUT

EMAIL COMES IN

ML-BASED SPAM CLASSIFIER

A LARGE BODY (CORPUS) OF SPAM AND HAM EMAILS

OUTPUT

SPAM OR HAM VERDICT

THE KEY DIFFERENCE BETWEEN THE ML AND THE RULE-BASED APPROACH IS:

THE ML-BASED APPROACH VARIES ITS ALGORITHM BASED ON WHAT THE DATA TELLS IT

WHILE THE RULE-BASED ALGORITHM IS STATIC

WHAT IS MACHINE LEARNING?

NOTE THAT WE MADE NO STATEMENT ABOUT WHICH IS MORE COMPLEX -

ITS ENTIRELY POSSIBLE THAT THE RULES-BASED APPROACH IS ACTUALLY FAR MORE COMPLEX THAN THE ML-BASED APPROACH

BUT THE DEFINING CHARACTERISTIC OF A MACHINE-LEARNING APPROACH IS THAT THE ALGORITHM ADJUSTS BASED ON DATA

WE SPOKE ABOUT HOW OUR ML-BASED SPAM DETECTOR COULD "LEARN FROM" A CORPUS OF DATA

(THE CORPUS : EMAILS EXPLICITLY MARKED BY USERS AS SPAM OR HAM)

WHAT ARE SOME OF THE WAYS IN WHICH THIS LEARNING COULD HAPPEN?

LET'S CYCLE THROUGH A FEW DIFFERENT TECHNIQUES