INTRO TO MACHINE LEARNING
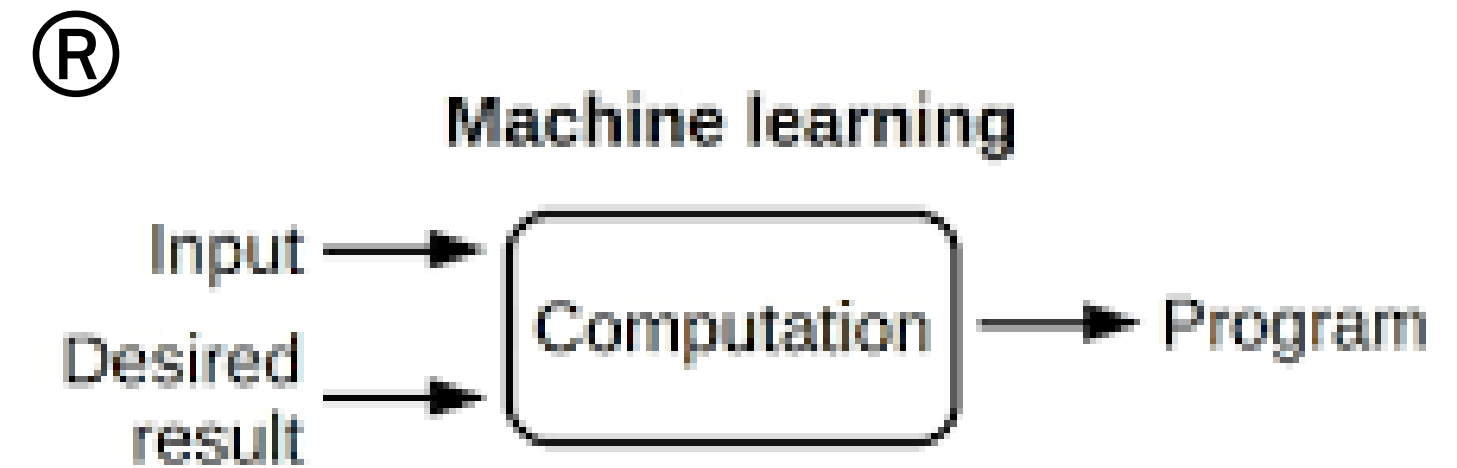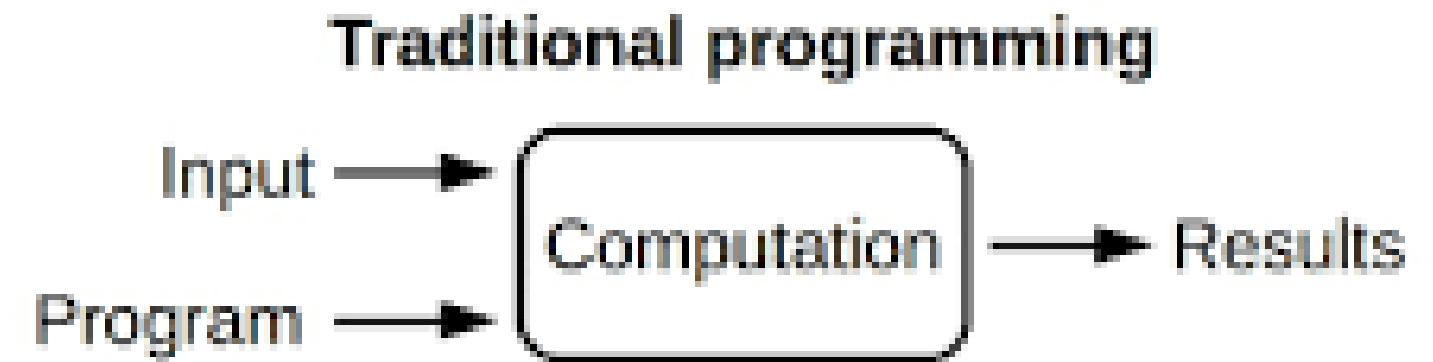
®

WHAT IS ML?

INTRO TO MACHINE LEARNING

WHAT IS ML?

Traditional programming

Input ⟶

Program ⟶ | Computation | ⟶ Results

®

Machine learning

Input ⟶

Desired result ⟶ | Computation | ⟶ Program

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

- Task T: what the model has to learn to perform
- Performance measure P: how well the model is doing
- Experience E: the dataset containing features with/without labels

Tom Mitchell

EXAMPLE

APPLICATIONS

CAN YOU THINK OF ANY?

E-COMMERCE

Recommendation

Use a Recommendation System

SIRI

Natural Language Preprocessing

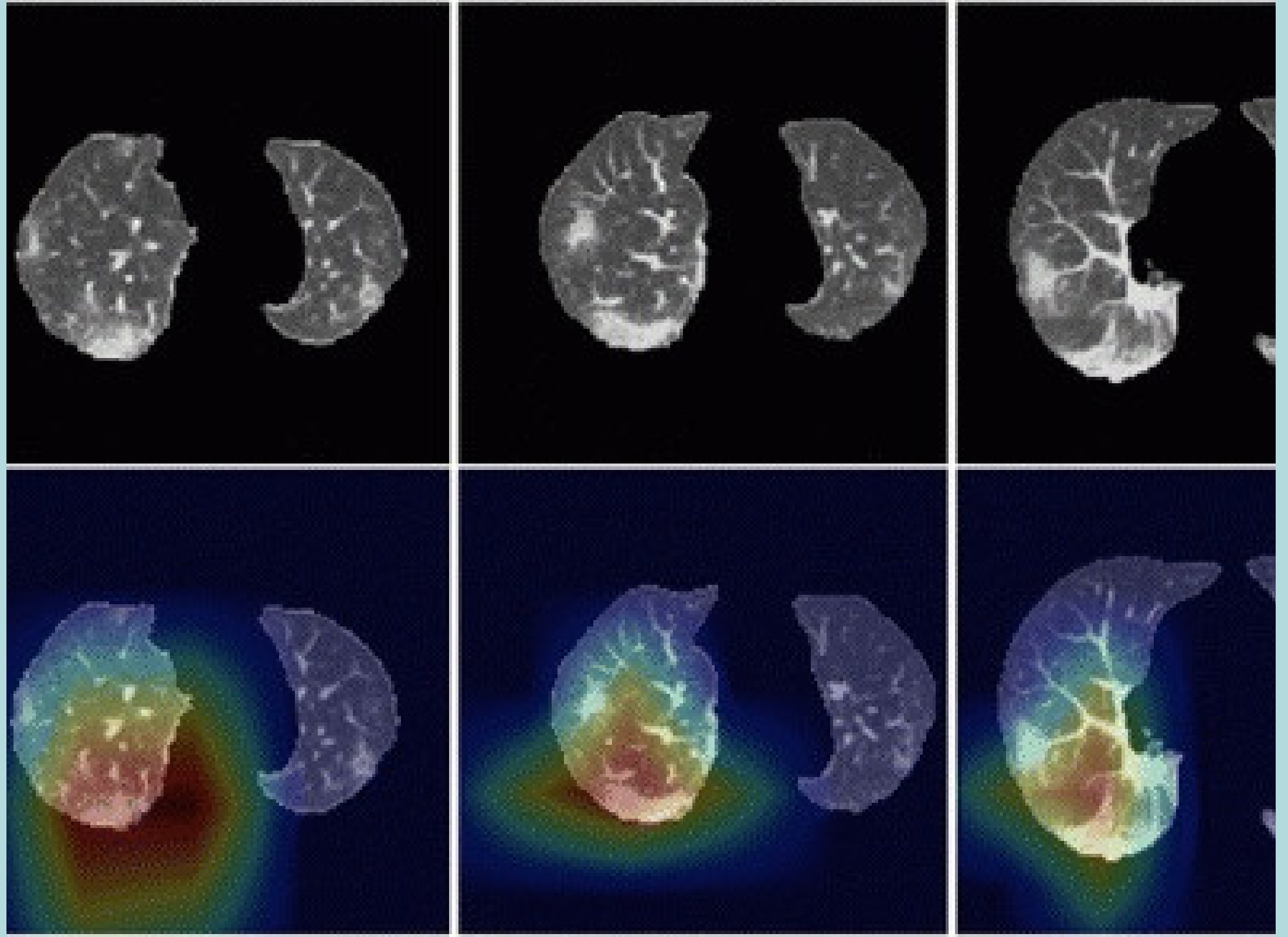**Answering questions; speech recognition; summarizing documents; classifying documents, text to speech**

MEDICINE

Computer Vision

Finding anomalies in radiology images, including CT, MRI, and X-ray images

Predict skin-cancer with near-human accuracy
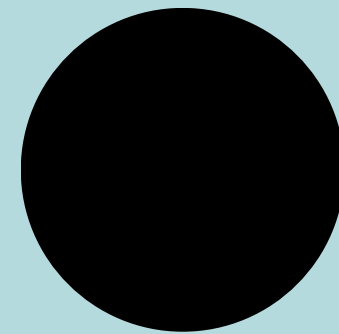
TYPES OF MACHINE LEARNING
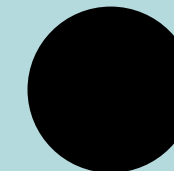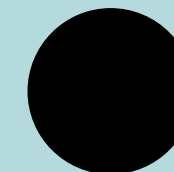
SUPERVISED LEARNING

CLASSIFICATION

REGRESSION

UNSUPERVISED LEARNING

CLUSTERING

ASSOCIATION

CASE STUDY

CAN YOU SURVIVE THE TITANIC?

Variable	Definition	Key
survival	Survival	0 = No, 1 = Yes
pclass	Ticket class	1 = 1st, 2 = 2nd, 3 = 3rd
sex	Sex	
Age	Age in years	
sibsp	# of siblings / spouses aboard the Titanic	
parch	# of parents / children aboard the Titanic	
ticket	Ticket number	
fare	Passenger fare	
cabin	Cabin number	
embarked	Port of Embarkation	C = Cherbourg, Q = Queenstown, S = Southampton
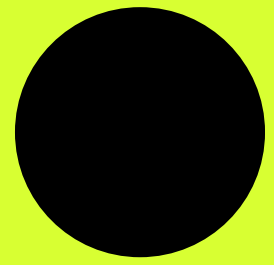
WHAT IS THE LABEL HERE?

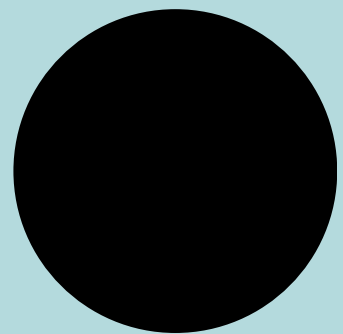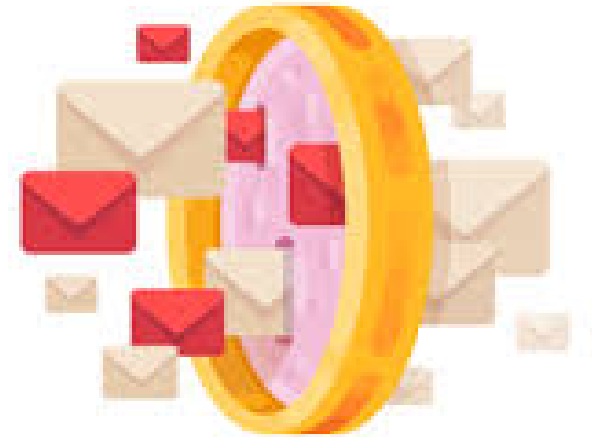|	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
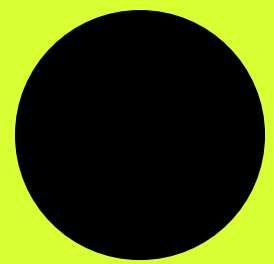
WHAT ARE THE INPUTS?
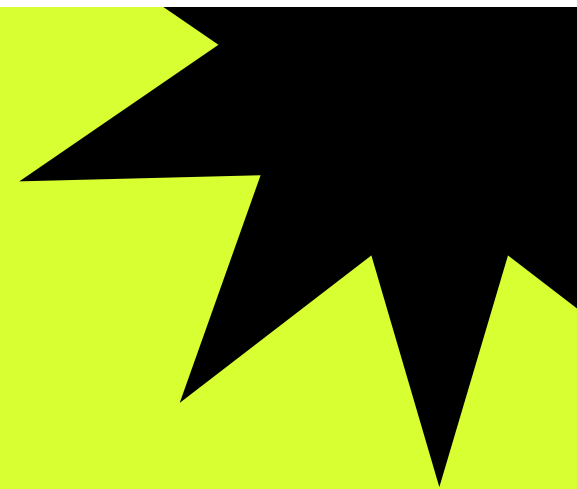
SUPERVISED LEARNING

CLASSIFICATION

FRAUD DETECTION

REGRESSION

TODAY TOMORROW