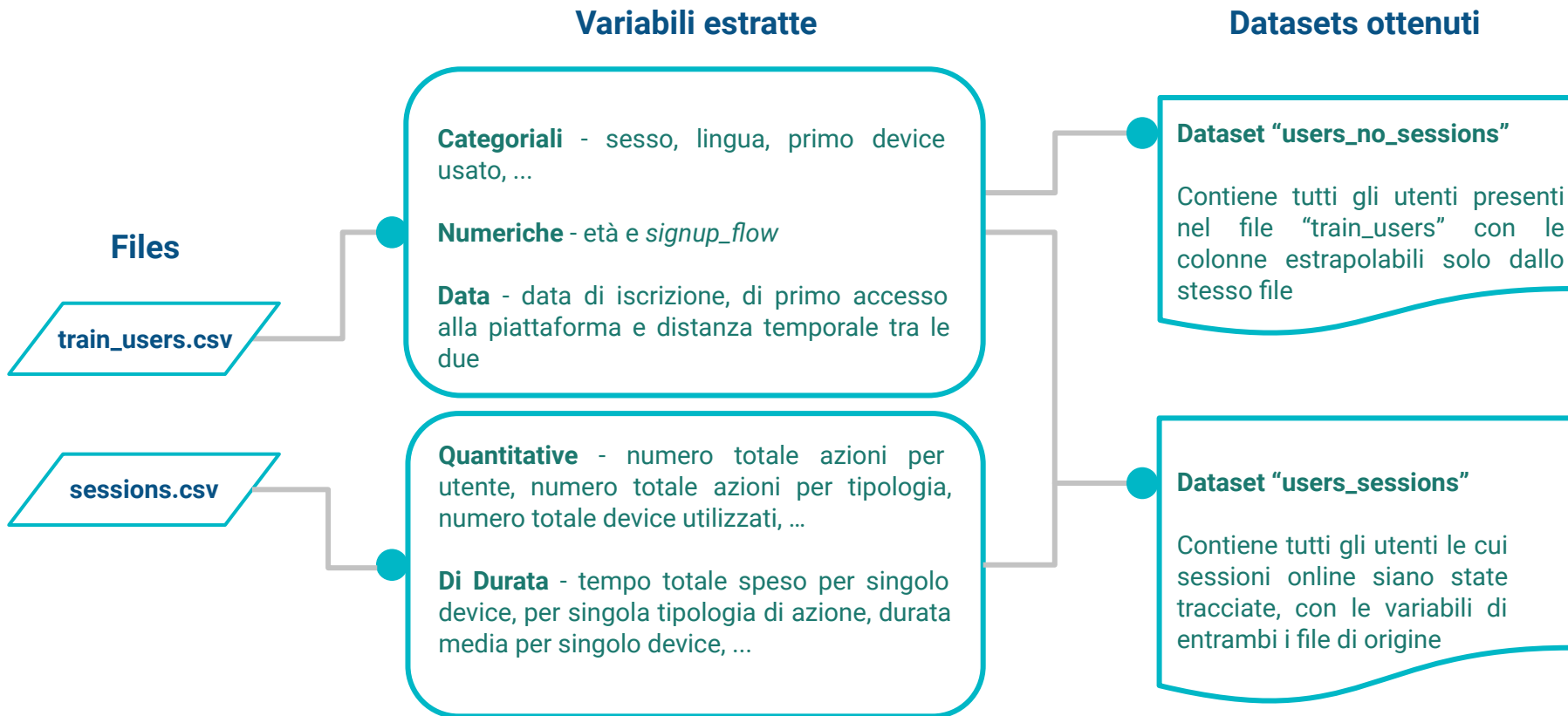
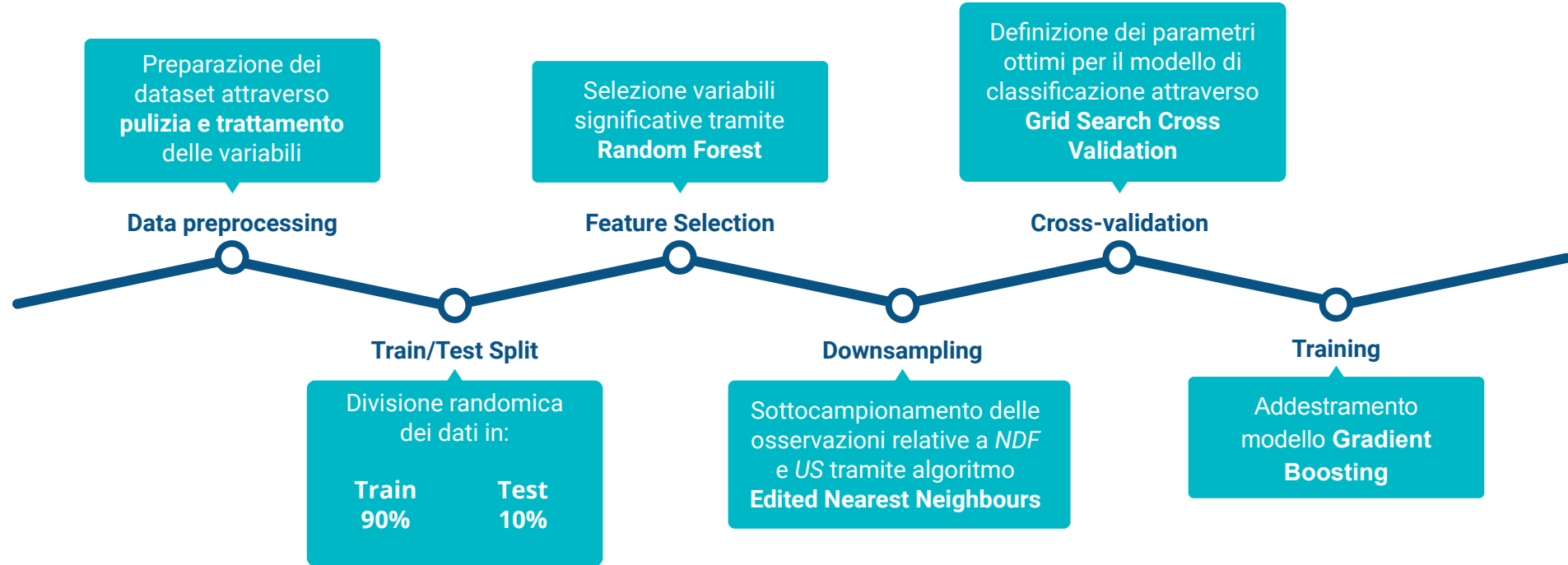

ConTe.it

First user booking

Creazione dei dataset

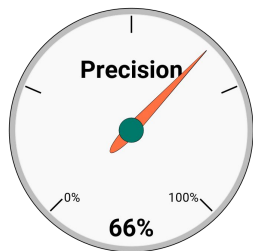


Approccio al problema

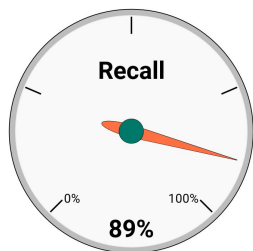


Risultati

Matrice di confusione del Test Set



Il modello indovina il **66%** delle sue previsioni NDF



Il modello intercetta l' **89%** delle osservazioni NDF reali

Reali

		Previsti							NDF		US		other	
reali	0	0	0	0	0	0	0	34	0	0	15	0		
	0	0	0	0	1	0	0	80	0	0	44	2		
	0	0	0	0	0	0	0	67	0	0	41	1		
	0	0	0	0	1	0	0	145	0	0	89	0		
	0	0	0	0	2	0	0	332	0	0	168	8		
	0	0	0	0	2	0	0	168	0	0	75	2		
	0	0	0	0	2	0	2	185	0	0	79	1		
NDF	0	0	0	0	7	0	1	11062	0	0	1324	40		
US other	0	0	0	0	0	0	0	50	0	0	32	1		
	0	0	0	0	0	0	0	12	0	0	7	1		
	1	0	0	0	7	0	0	3871	0	0	2302	55		
	0	0	0	1	2	0	0	667	0	0	336	21		

Risultati

User	P1	P2	P3	P4	True
q9p###	US	other	NDF	GB	other
9bz###	NDF	US	other	GB	US
kux###	NDF	US	other	AU	US
qqn###	NDF	US	other	FR	NDF
nw9###	NDF	US	other	FR	NDF
oop###	NDF	US	other	FR	IT
a62###	NDF	other	US	FR	NDF
611###	NDF	other	US	FR	FR
qbs###	NDF	US	FR	other	US
f2k###	NDF	US	other	FR	US
...

Ranking

La tabella contiene:



User - nome utente (parzialmente oscurato)



P1_P4 - le prime 4 previsioni per singolo utente in ordine di probabilità, partendo dalla più probabile (P1)



True - la destinazione effettiva

Nel campione di test estratto a sinistra, per **9 utenti su 10** il modello attribuisce una buona probabilità alla destinazione effettiva. Nella popolazione totale di test, tale rapporto si attesta a **9,4 su 10**.

Il modello, quindi, si dimostra essere in grado di fornire una **lista accurata** di (poche) possibili destinazioni.

Conclusioni

Summary

Previsione

Il modello addestrato non si è dimostrato in grado di distinguere tra le osservazioni più frequenti (*NDF*, *US* e *other*) e gli altri Paesi di destinazione...

Ranking

Tuttavia, quando si considera il **ranking** attribuito dal modello in base alle probabilità, osserviamo come lo stesso sia in grado di fornire una lista molto accurata di possibili destinazioni.

Linee guida

Se l'obiettivo è una campagna di marketing mirata, **si suggerisce di focalizzare sforzi e risorse verso i primi 4 Paesi previsti per ogni utente.**

Possibili miglioramenti

In fase di analisi è sembrato evidente come l'impossibilità di ottenere le informazioni delle attività online di ogni utente prima del 2014 sia stato un grosso limite per l'addestramento. Se tale limitatezza è dettata dall'esigenza di contenere le dimensioni dei file, si potrebbe contenere la selezione degli utenti ad un solo anno ma rendere disponibili le informazioni sulle sessioni online per tutti.