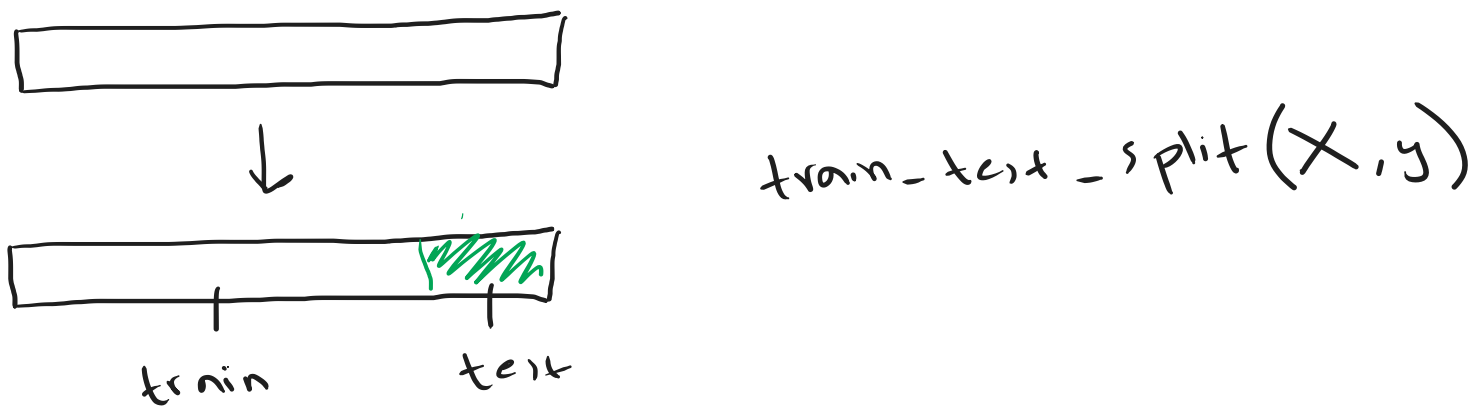


Train/Validation/Test split

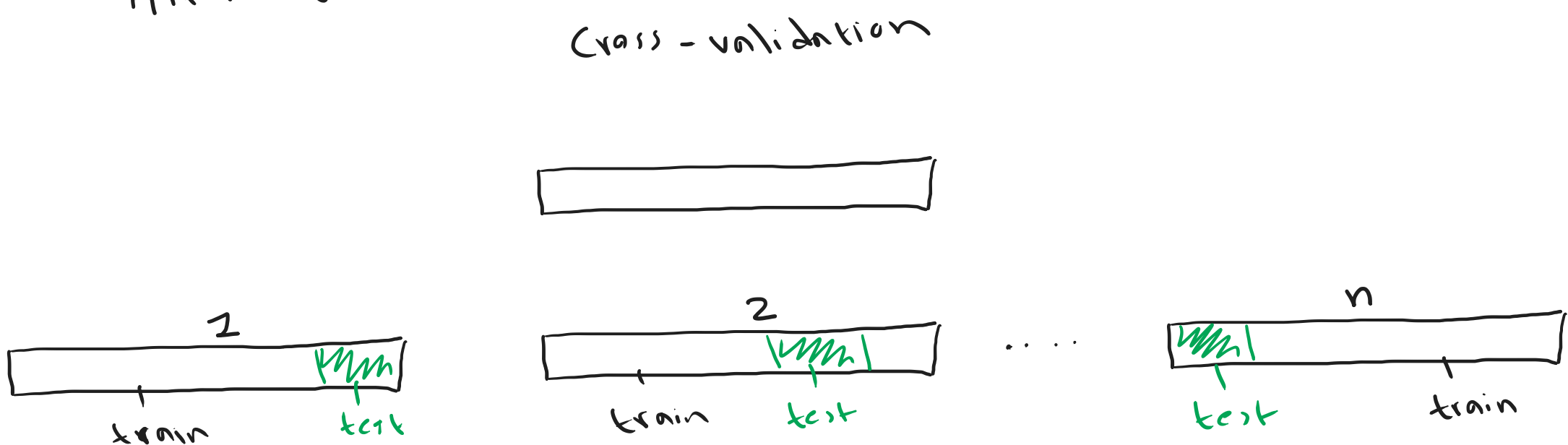
Vi har tidigare lärt oss om vikten av att korrekt evaluera modeller. En huvudsats är att vi måste säkerställa att vi inte evaluerar på den datan vi tränar - eftersom att vi vill simulera prestanda på verklig data, och den datan är ju osedd.

Vi har lärt oss två olika metoder tidigare, i växande sofistikeringsgrad.

Alternativ 1:

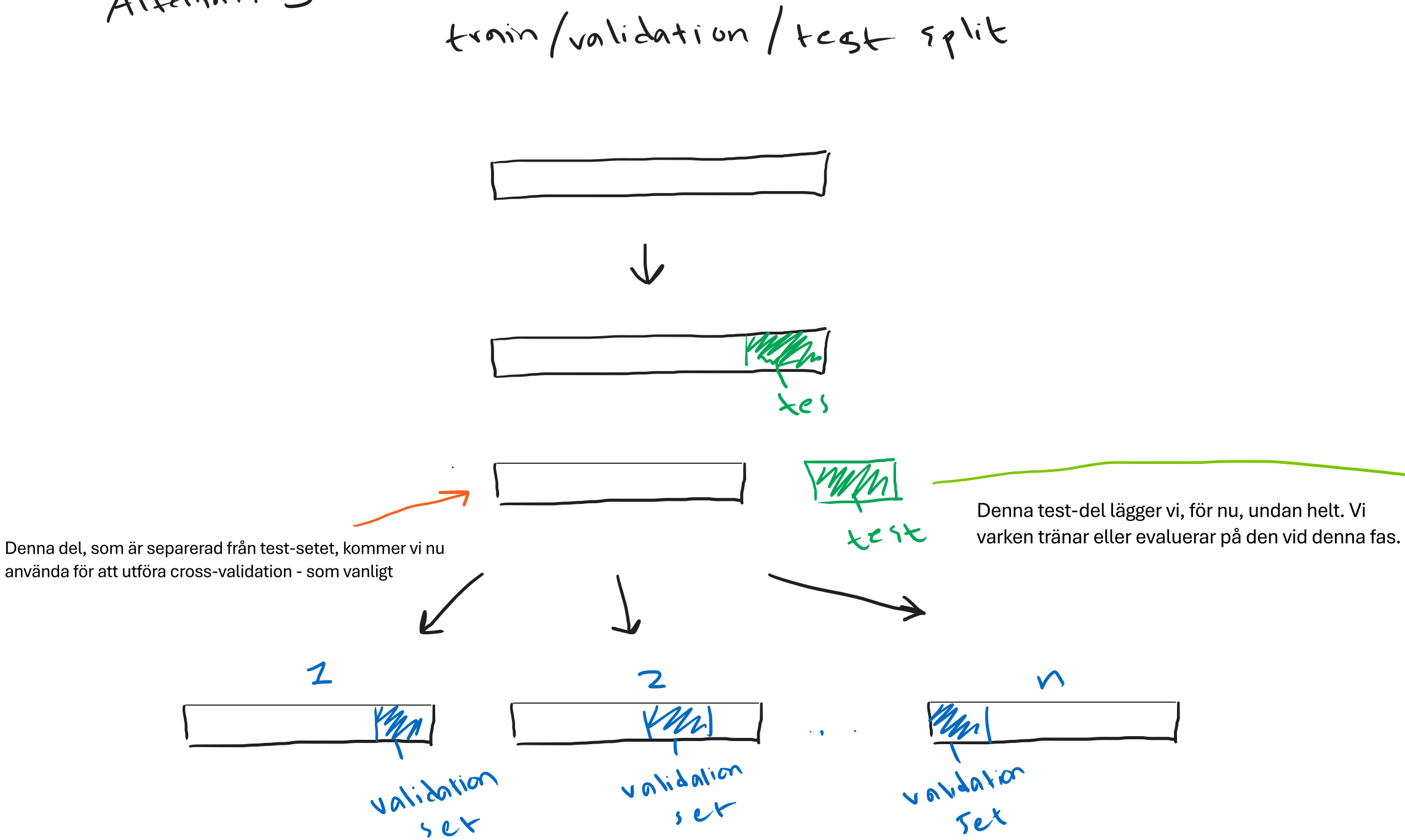


Alternativ 2:



Styrkan med CV är att den mitigerar problemet med att man "får otur" ifall man använder alternativ 1 och får ett test-set som inte är representativt - dvs, där din ex overfittade model prestanda på produktions data blir överskattad & således missvissande.

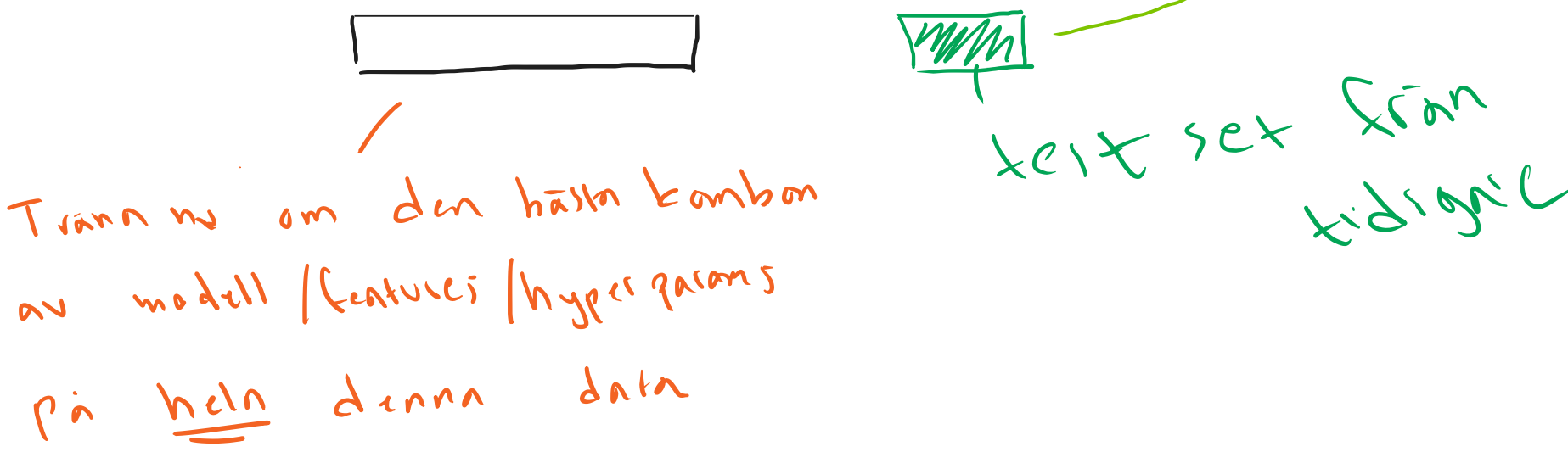
Alternativ 3:



Så, vi lägger alltså undan (omedelbart innan någon träning skett öht) en del av datan (test-set).

Övrig data använder vi för att utföra cross-validation på. Vi använder alltså denna data för att hitta bästa features och modell, samt för att utföra hyperparameter tuning på.

När vi har hittat den modell/features/hyperparameter-kombination som presterar bäst på denna träningsdatan, gör vi följande



Vi evaluerar sedan på vårt test set!

Resultatet från evalueringen av ditt test-set är vad du sedan rapporterar in som slutgiltigt resultat!

