

PROVJERA (VALIDACIJA) MODELA

POGLAVLJE 18.4

Prema slajdovima Stuarta Russela, Krunoslava Puljića, Tomislava Šmuca, Hantao Zhanga (hvala)!

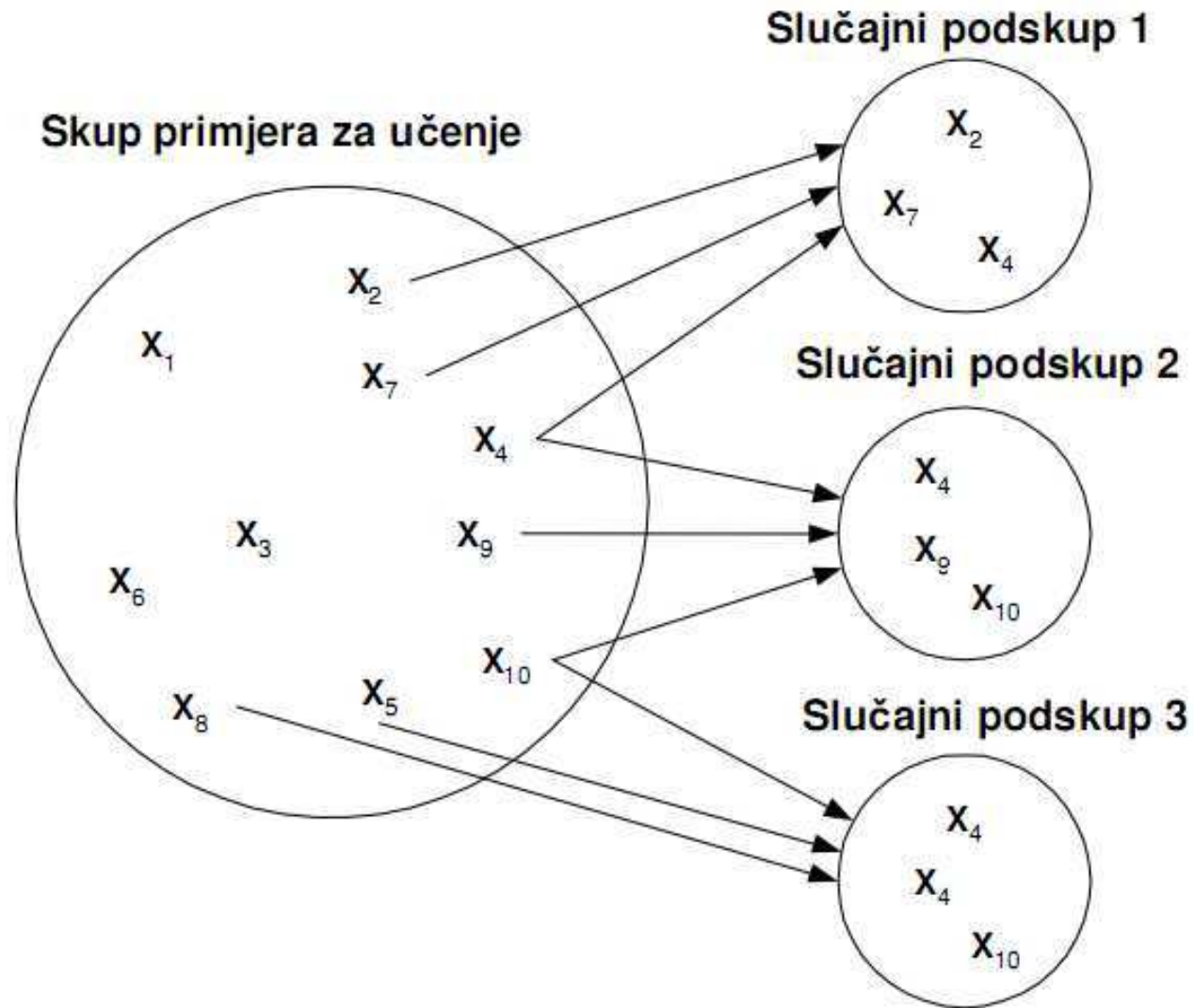
Poželjni modeli

- u svakom slučaju trebao bi imati minimalan broj grešaka $\min e(h, T)$
 - u idealnom slučaju $h(x_i) = y_i$ (interpolacija).
- ne bi smio biti previše složen
 - za neku mjeru $C(h)$ složenosti modela.
- trebaju nam procedure za pretraživanje prostora mogućih modela ($h \in H$) koje vode računa i o greškama $e(h)$ i o složenosti $C(h)$:
 - $\min(e(h, T) + \alpha C(h))$ “regularizacija”.

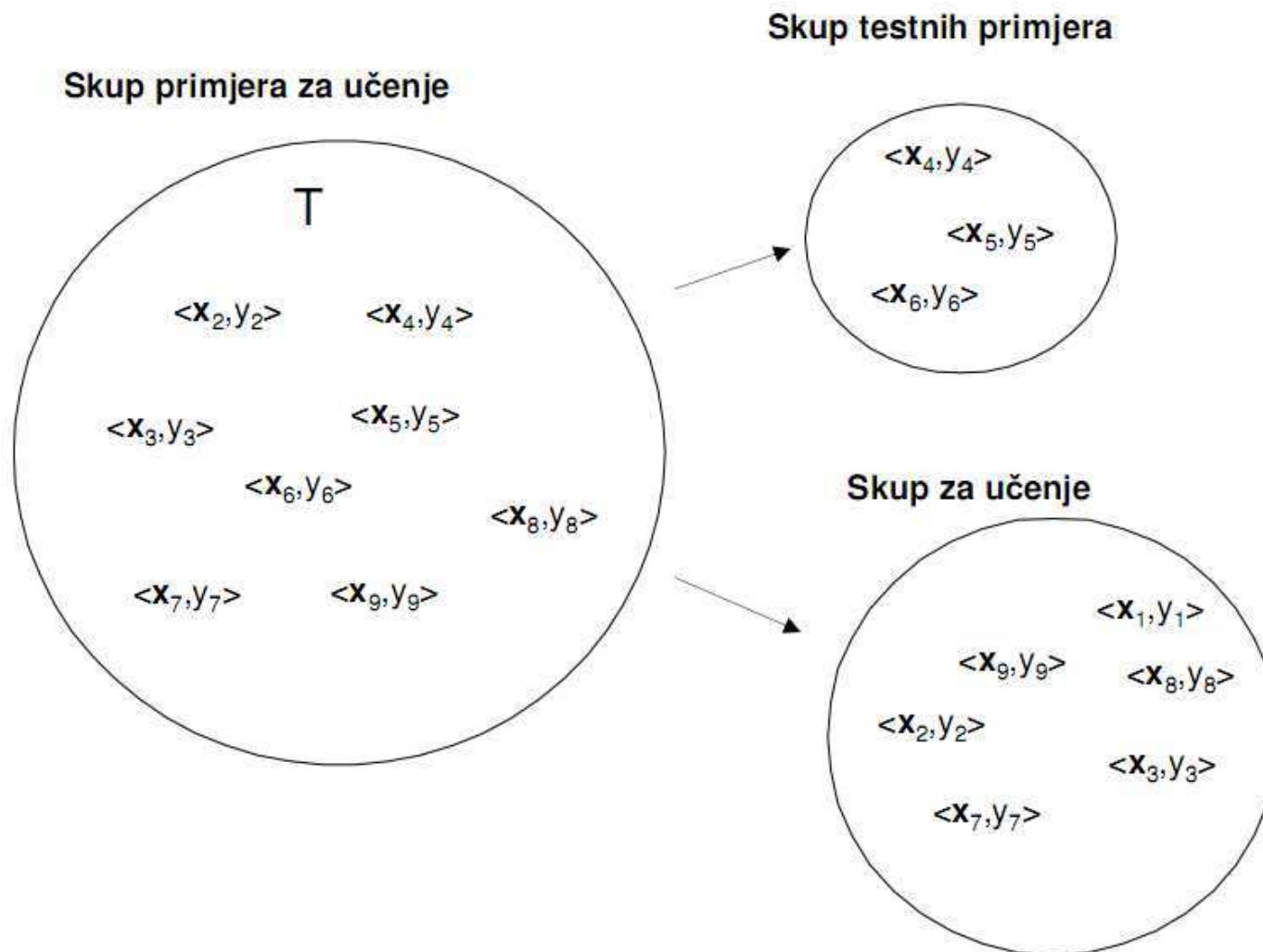
Validacija modela

- Procjena stvarne greške modela
- Tehnike probira (engl. Sampling—Resampling):
 - Train & Test metoda
 - Unakrsna validacija (engl. Cross—Validation)
 - Pojedinačna unakrsna validacija (LOOCV)

Tehnike probira



Train & test metoda



Train & test metoda

- **Slučajno** odaberemo $1/3$ od dostupnih primjera za učenje i stavimo ih u novi skup — skup za testiranje (engl. Test set).
- Ostatak od $2/3$ primjera iskoristimo za učenje modela — skup za učenje (engl. Training set).
- Naučimo model na skupu za učenje.
- Procijenimo stvarnu grešku modela “testirajući” novi model na skupu za testiranje.

Train & test metoda

Svojstva:

– dobro:

- Jednostavna metoda — odabiremo onaj model koji daje najmanju grešku na testnom skupu.

– loše:

- Gubimo vrijedne podatke — $1/3$ podataka se uopće ne koristi za izradu modela.
- Ako imamo relativno malo podataka za učenje, ocjena greške na testnom skupu je vrlo nepouzdana (\gg varijanca greške).

K-struka unakrsna validacija

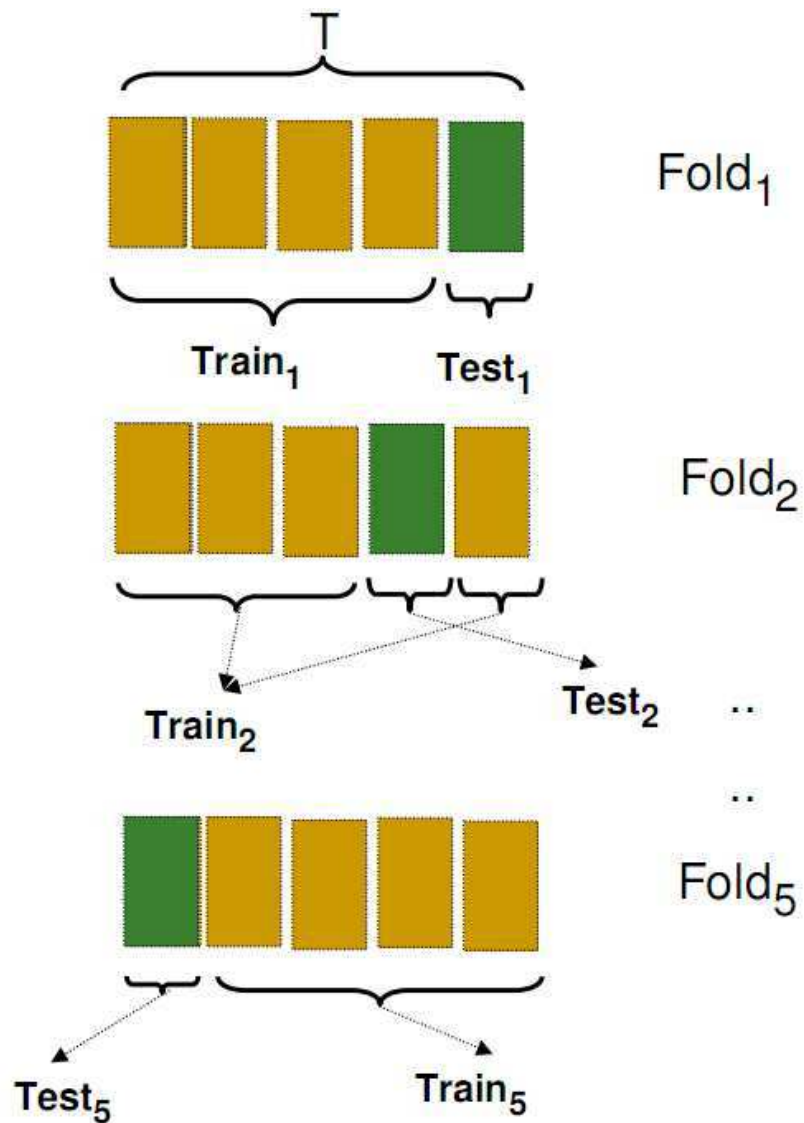
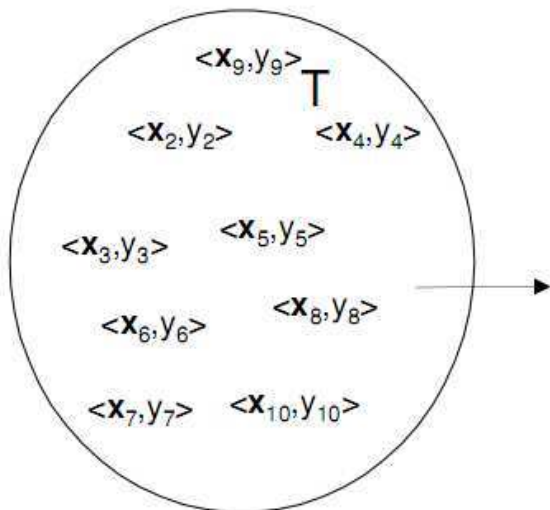
K-struka unakrsna validacija (engl. k-fold cross validation)

- Slučajno rasporediti primjere za učenje u k odvojenih skupova T_i , za $i = 1, \dots, k$,
— tipično po 30+ primjera.
- Za $i = 1$ do k ponovi:
 - Koristi T_i kao **testni** skup, a ostale podatke T_m , $m \neq i$, iskoristi za **učenje** modela h_i .
 - Na testnom skupu T_i izračunaj grešku L_i modela h_i .
- Izračunaj prosječnu grešku za svih k modela

$$L = \sum_{i=1}^k L_i / k.$$

K-struka unakrsna validacija

Skup primjera za učenje



Pojedinačna unakrsna validacija

Pojedinačna unakrsna validacija (LOOCV)
(engl. Leave-one-out cross validation)

- Na skupu primjera za učenje $(x_i, y_i) \in D, i = 1, \dots, N$
- Za $i = 1$ do N ponovi:
 - Privremeno izdvoji primjer (x_i, y_i) iz skupa primjera za učenje.
 - Nauči model h_i na preostalim $N - 1$ primjera.
 - Izračunaj grešku L_i modela h_i na primjeru (x_i, y_i) .
- Izračunaj prosječnu grešku za svih N modela

$$L = \sum_{i=1}^N L_i / N.$$

Svojstva evaluacijskih metoda

Metoda	Dobro	Loše
Train & Test	<ul style="list-style-type: none"> • Jeftina – učimo samo jednom 	<ul style="list-style-type: none"> • Gubimo puno primjera za učenje • Nepouzdana procjena stvarne greške
K-fold CV	<ul style="list-style-type: none"> • Gubimo samo N/k za učenje jednog modela • Stabilnija procjena greške 	<ul style="list-style-type: none"> • K puta “skuplja” od T & T • Učimo k modela
LOO CV	<ul style="list-style-type: none"> • Praktički učimo na svim primjerima – 1 • Dobra za mali broj primjera 	<ul style="list-style-type: none"> • Vrlo skupa za veliki N • Učimo N modela