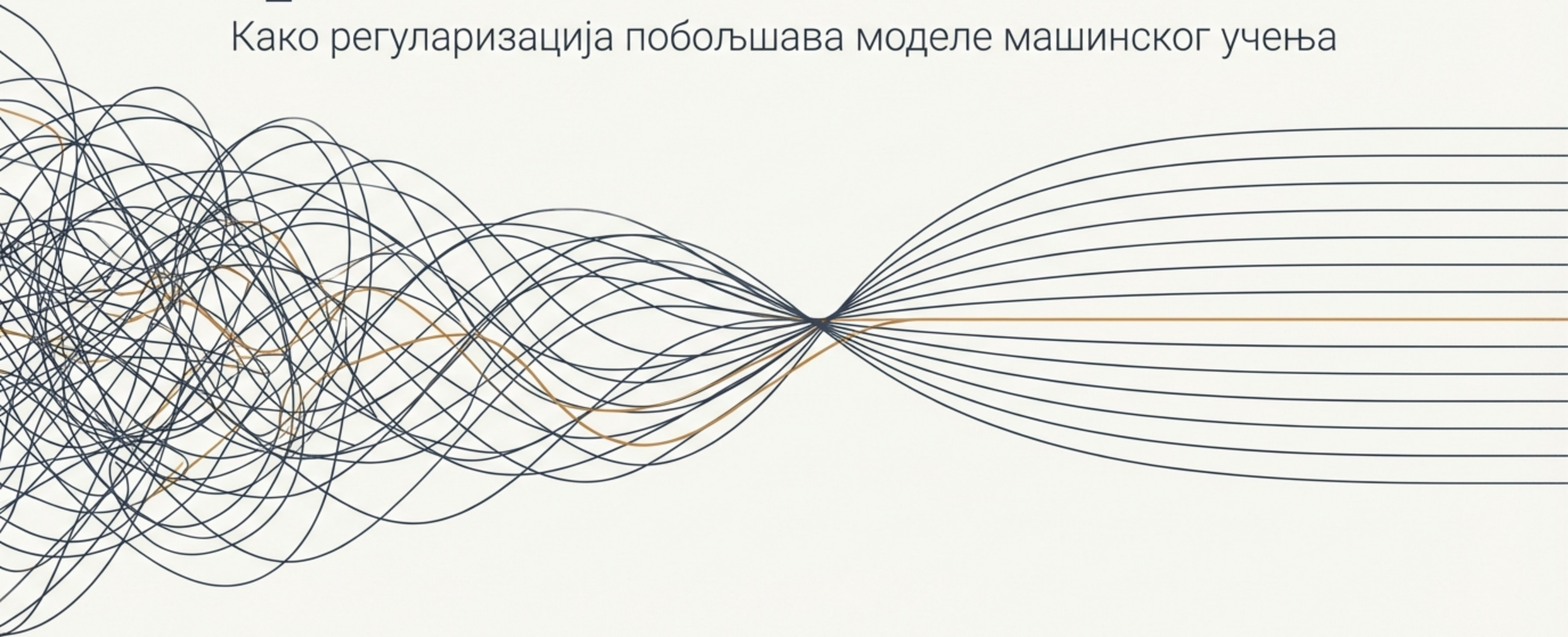


Кроћење комплексности

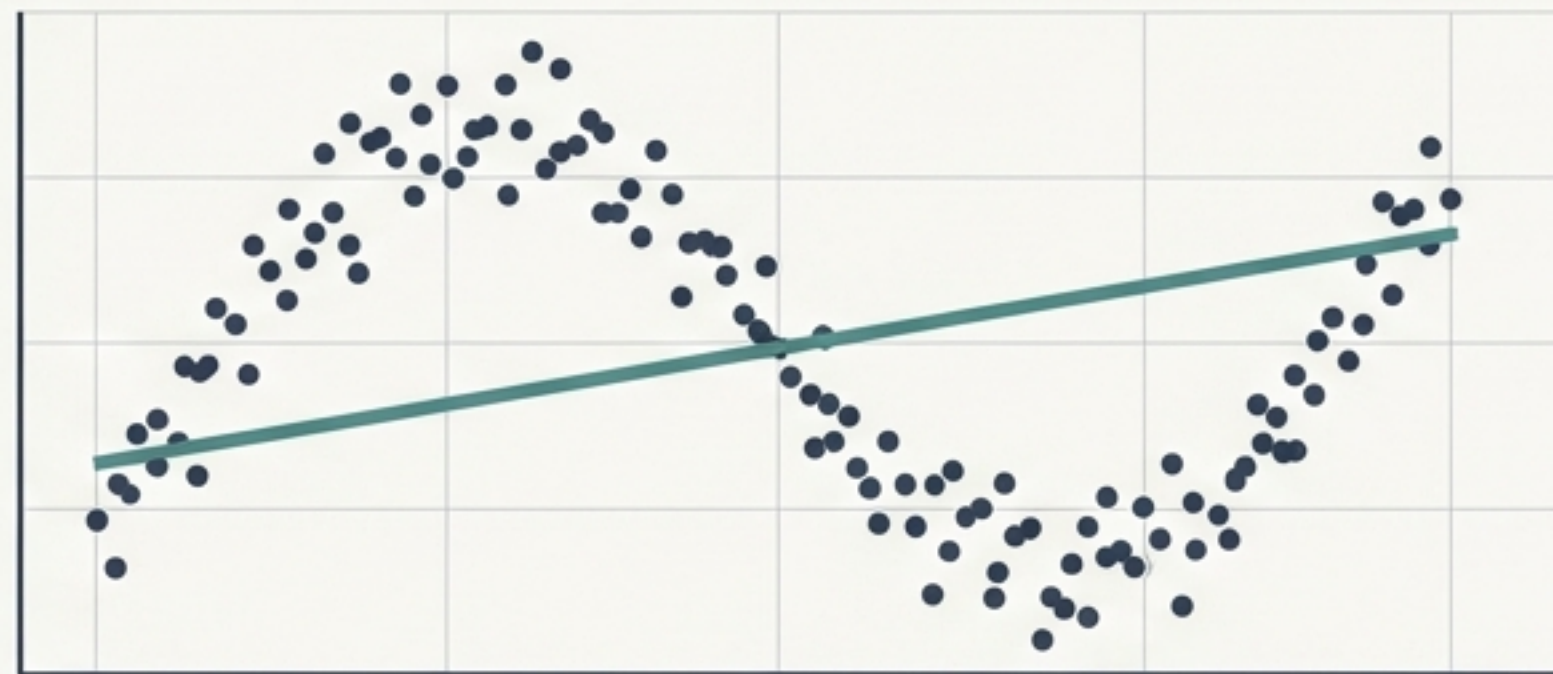
Како регуларизација побољшава modele машинског учења



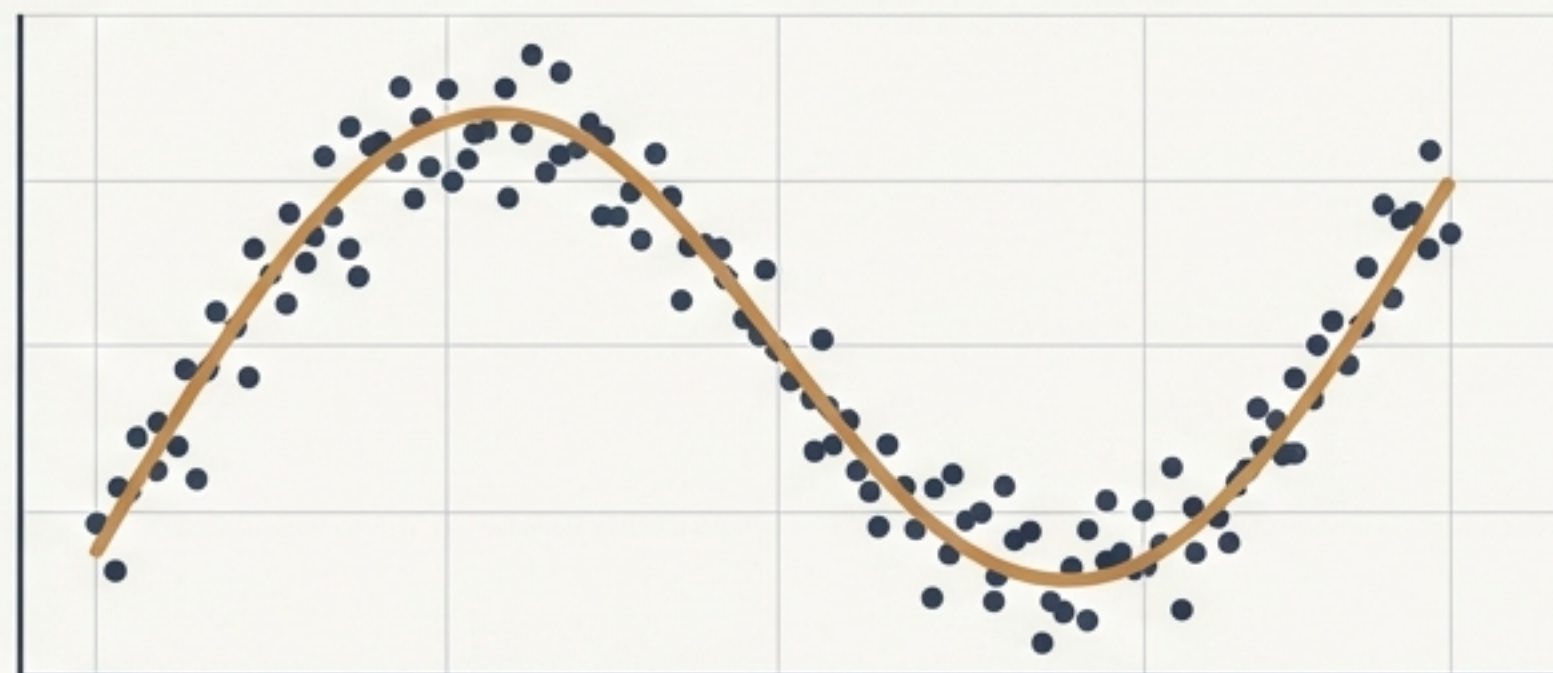
Почињемо са циљем: моћнији модели.

У машинском учењу, тежимо ка креирању комплексних модела. Зашто? Јер нам они помажу да научимо богатији скуп зависности у подацима, откривајући суптилне обрасце које једноставнији модели не могу да виде. Већа комплексност обећава већу прецизност.

Једноставан модел



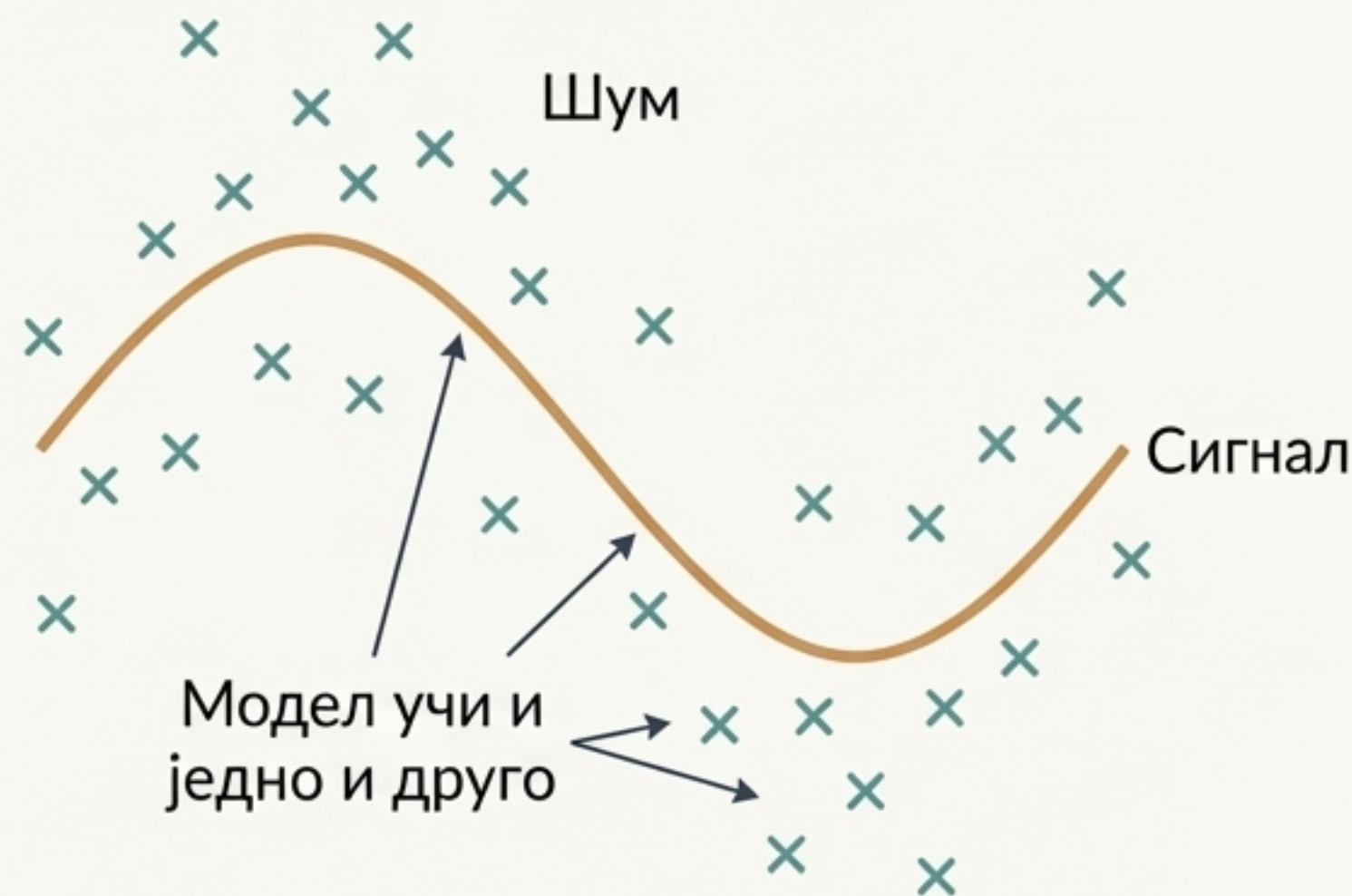
Комплексан модел



Али велика моћ носи скривени ризик.

Проблем се назива **преприлагођавање (overfitting)**. Модел постаје толико добар у учењу података за тренирање да почиње да “памти” њихов шум и случајне флуктуације, уместо да учи само основни сигнал.

Замислите га као студента који је напамет научио уџбеник, али не разуме материју довољно да реши нови, непознати проблем.



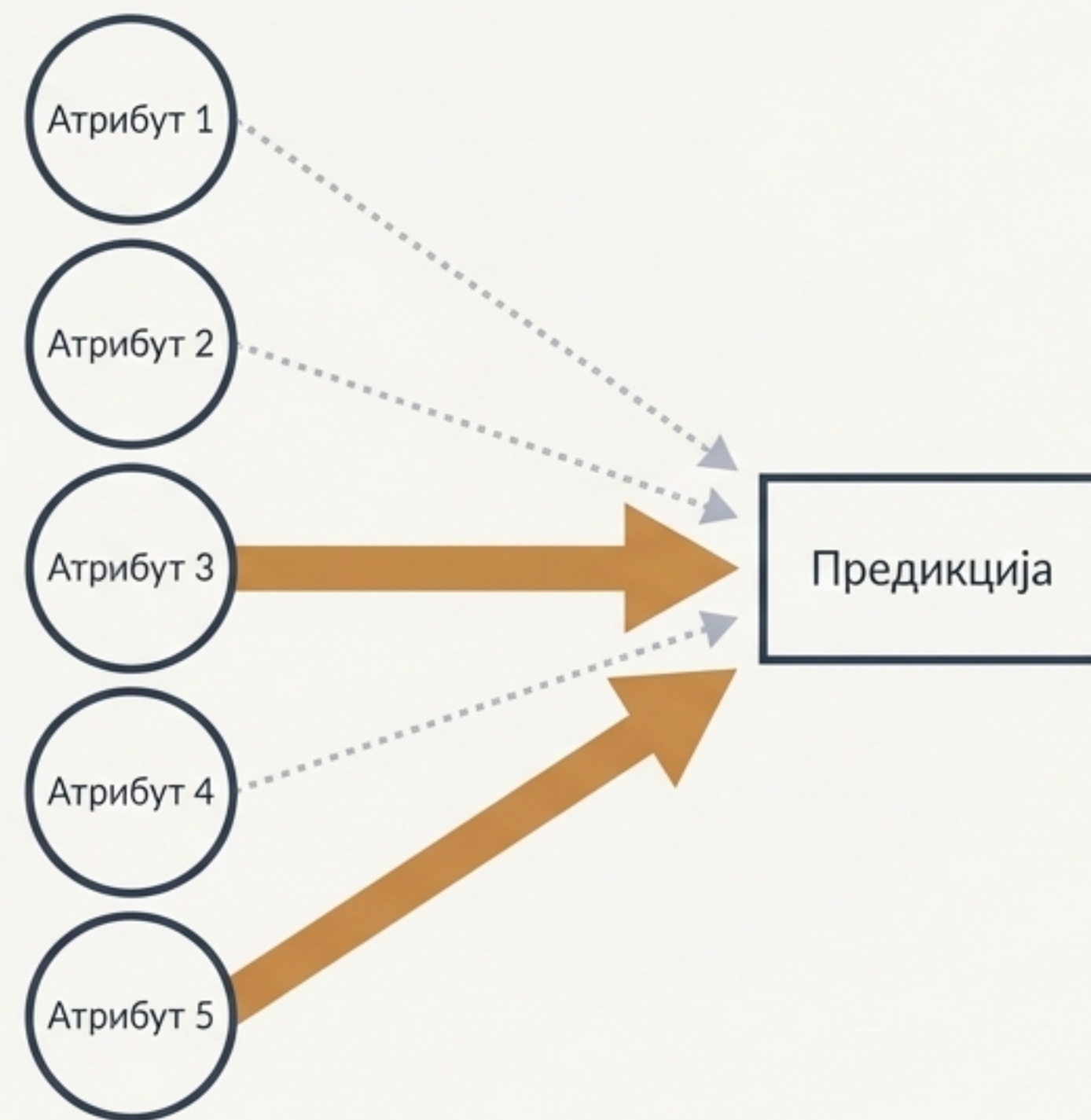
Јасан симптом: неколико параметара доминира.

Када обучимо модел и погледамо вредности његових параметара, често можемо уочити знак преприлагођавања. Неколико параметара постаје изузетно велико у поређењу са осталима. Ови "доминантни" параметри имају несразмерно велик утицај на коначну предикцију.



Последица: утицај осталих атрибута се занемарује.

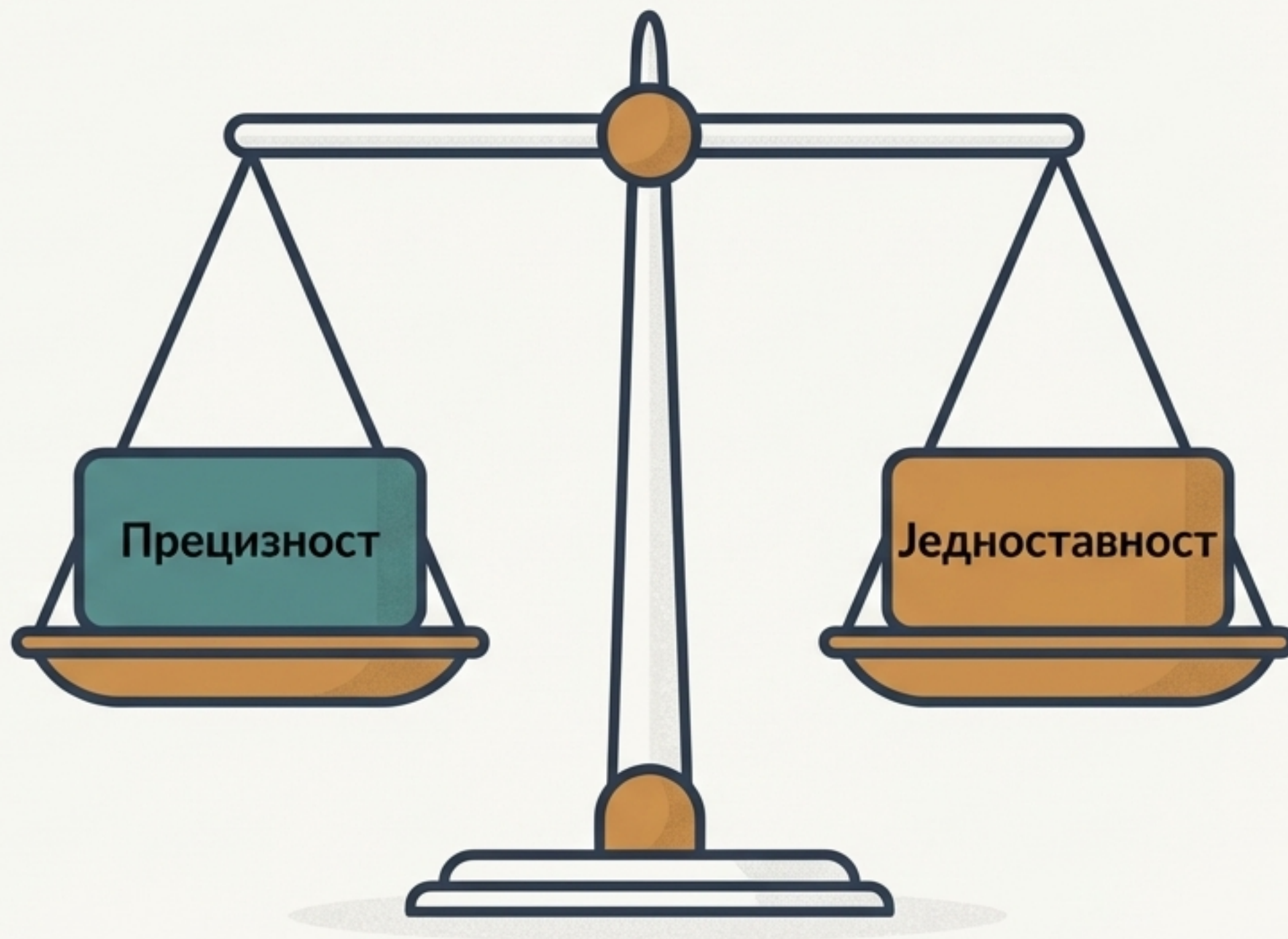
Ови атрибути са великим параметрима (у нашем примеру, атрибути 3 и 5) могу да **занемаре утицај преосталих атрибута** на вредност предикција. Модел постаје превише осетљив на мале промене у овим доминантним атрибутима, док игнорише корисне информације из осталих. Ово понашање је јасан вид преприлагођавања подацима.



Наш нови задатак: пронаћи равнотежу.

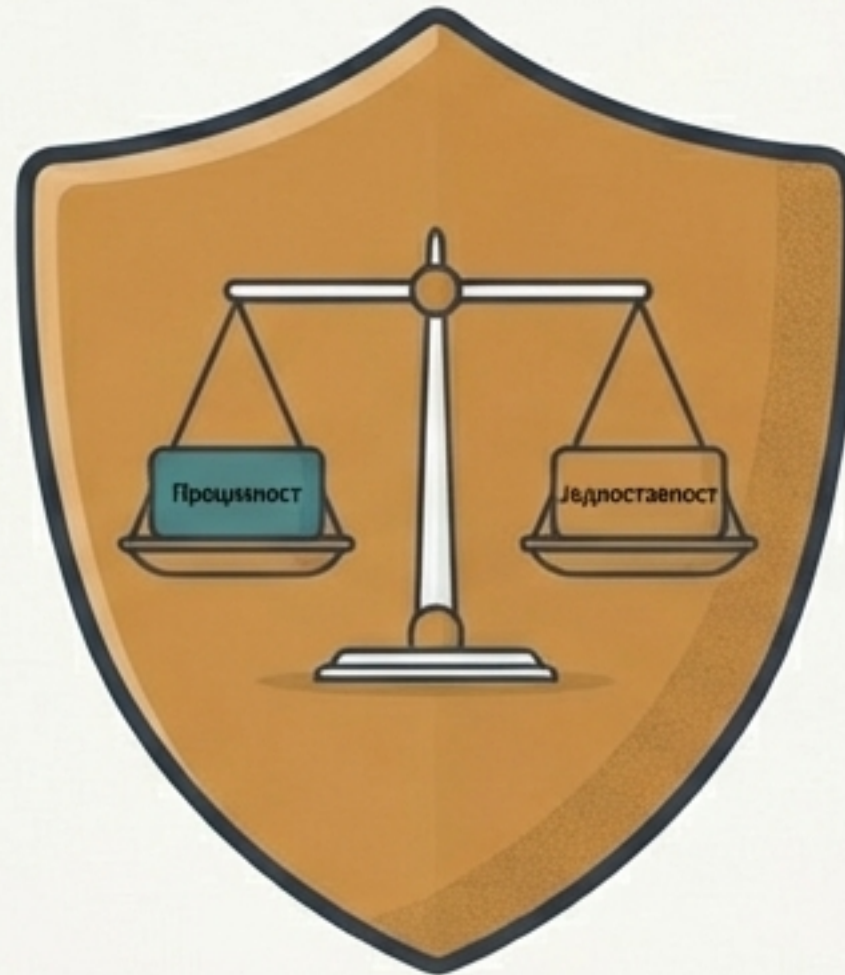
Циљ није да елиминишемо комплексне моделе, већ да их контролишемо. Пожељно је, у некој мери, **ограничити вредности параметара**.

Желимо да модел научи параметре који одражавају својства података, али истовремено желимо да пратимо њихову величину како бисмо спречили преприлагођавање.



Решење се зове **Регуларизација**.

Регуларизација (*eng. regularisation*) је скуп техника које додају “казну” (*penalty*) за велике вредности параметара у процес учења модела. Уместо да модел оптимизује само своју грешку на подацима, сада **мора да балансира између тачности и величине својих параметара**.



Како то функционеше? Корак 1: Полазна тачка.

У стандардној линеарној регресији, циљ је минимизовати средњеквадратну грешку (Mean Squared Error). Ово је наш први и једини циљ: бити што тачнији на подацима за тренирање.

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Циљ 1: Бити прецизан

Корак 2: Корак 2: Додајемо нови циљ.

Регуларизација уводи нови члан у функцију грешке. Овај члан мери 'величину' модела сабирањем квадрата свих његових параметара. Ово постаје наш други циљ: одржати параметре малима.

$$\sum \theta_j^2$$

Циљ 2: Бити једноставан

Корак 3: Комбинујемо циљеве.

Сада комбинујемо оба циља у једну функцију грешке. Модел мора истовремено да минимизује грешку на подацима и да одржава параметре малима.

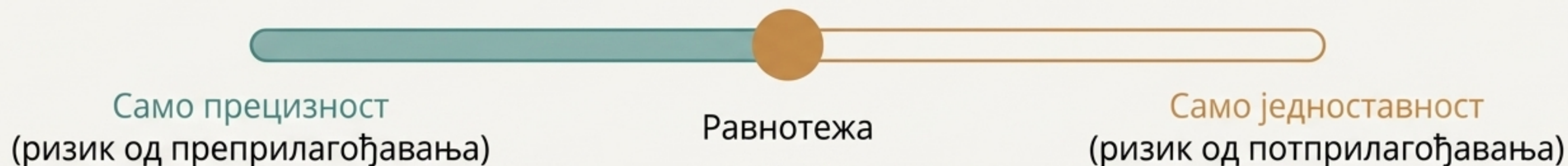
$$J(\theta) = \left[\frac{1}{2m} \sum (h_{\theta}(\mathbf{x}^{(i)}) - y^{(i)})^2 \right] + \left[\lambda \cdot \sum \theta_j^2 \right]$$

Прецизност

Казна за комплексност

Контролно дугме: Хиперпараметар λ (ламбда).

Вредност λ је хиперпараметар којим утичемо на **јачину регуларизације**. Она одређује колико је важан циљ 'бити једноставан' у односу на циљ "бити прецизан".



- **Ако је $\lambda = 0$:** Регулација нема никаквог ефекта. Враћамо се на стандардну линеарну регресију.
- **Како λ расте:** Казна за велике параметре постаје све јача, приморавајући модел да буде једноставнији, чак и по цену мање прецизности на тренинг скупу.

Зашто баш квадрати?

- Употреба квадрата у регуларизационом члану је намеран избор из два техничка разлога:
 1. **Спречава поништавање:** Квадрирањем се осигурава да су сви чланови позитивни. Тиме се спречава да се велики позитивни и велики негативни коефицијенти међусобно потиру.
 2. **Олакшава оптимизацију:** Овај облик функције грешке чува важна математичка својства (као што је конвексност) која омогућавају ефикасну примену алгоритама за оптимизацију, попут градијентног спуста.



Овај приступ има своје име.

Овако проширен облик линеарне регресије, допуњен регуларизационим чланом који користи суму квадрата параметара, назива се:

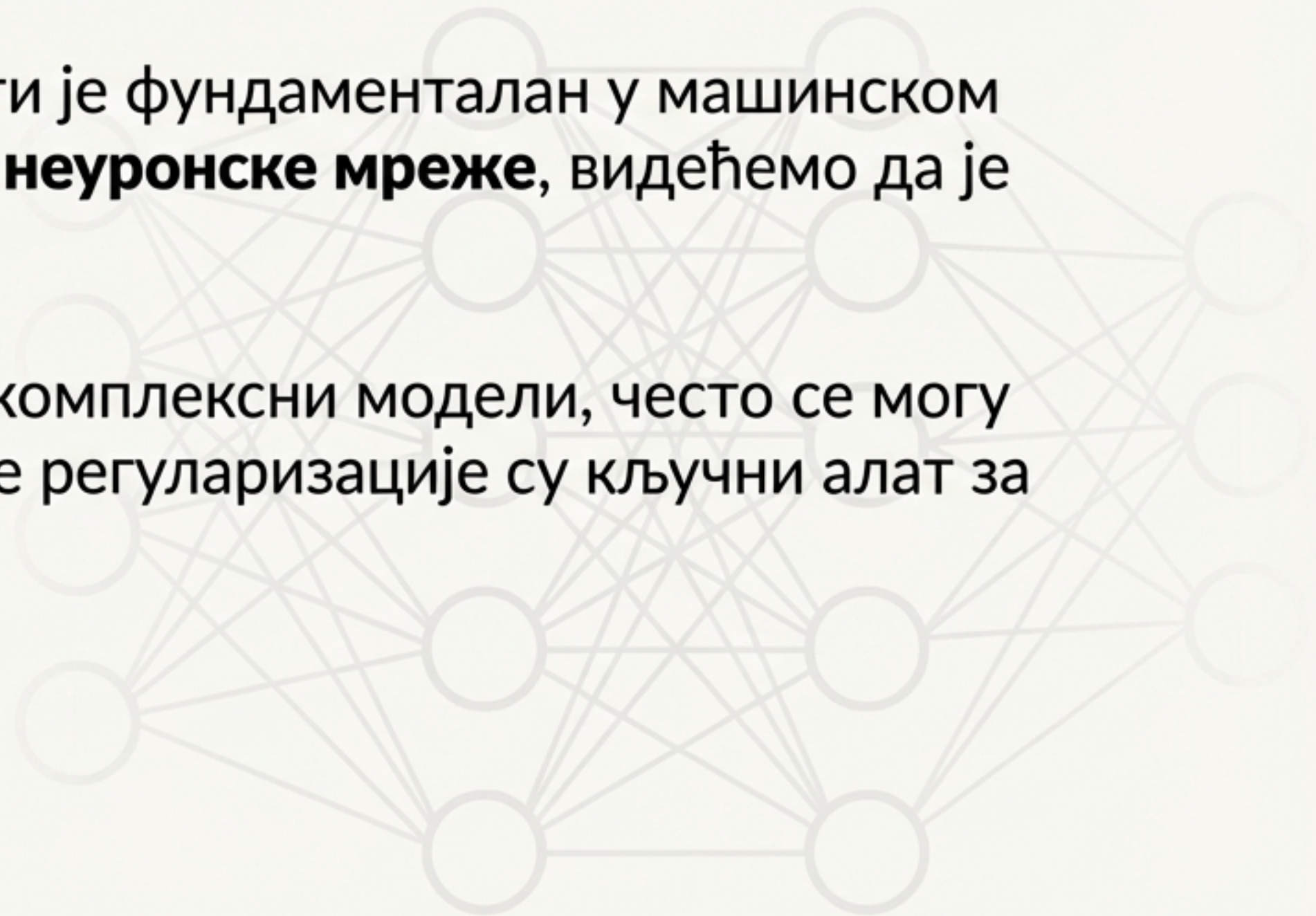
Гребена линеарна регресија

(eng. *Ridge Regression*)

Ово је тек почетак приче о регуларизацији.

Концепт кажњавања комплексности је фундаменталан у машинском учењу. Касније, када будемо увели **неуронске мреже**, видећемо да је регуларизација још важнија.

Пошто су неуронске мреже веома комплексни модели, често се могу преприлагодити подацима. Технике регуларизације су кључни алат за њихову контролу и обуку.



Кључне идеје за памћење.



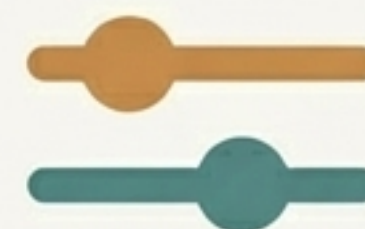
Проблем

Комплексни модели ризикују **преприлагођавање**, што се често манифестује кроз неколико екстремно великих параметара.



Решење

Регуларизација додаје 'казну' за велике параметре у функцију грешке, приморавајући модел да балансира прецизност и једноставност.



Алат

Гребена регресија користи суму квадрата параметара као казну, чија се јачина контролише хиперпараметром λ .