

# Učící se algoritmy aplikované na předvídání vzniku plazmatu v tokamaku

Helena Valouchová – Gymnázium Brno, m.h.valouchova@seznam.cz  
Vojtěch Stránský – Gymnázium Teplice,  
vojtech.stransky@post.gymtce.cz

## Abstrakt:

Cílem tohoto projektu bylo statisticky předvídát vznik plazmatu v tokamaku pomocí učících se algoritmů. Nejdříve jsme do programu nahráli databázi údajů z pokusů z předchozích let, ten poté předvídál výsledky aktuálních „výstřelů“. Nakonec jsme vyzkoušeli takové parametry, u kterých program odhadoval pravděpodobnost okolo hraniční hodnoty 50%.

## Může počítač „věštit“ výsledky pokusů?

Učící se algoritmy (SVM – Support Vector Machine) jsou metoda, která se z předchozích výsledků daného jevu snaží předpovídat výsledky budoucí. K tomu je zapotřebí velké množství dat (množství roste s počtem parametrů) a je nutno rozdělit výsledky na pozitivní k danému jevu a negativní. Tuto metodu lze tedy použít pro jevy, kdy lze rozlišit konečné množství tříd událostí.

V našem miniprojektu jsme aplikovali učící se algoritmy na pokusy v tokamaku GOLEM a zkoumali jsme pravděpodobnost vzniku plazmatu pro zadané parametry.

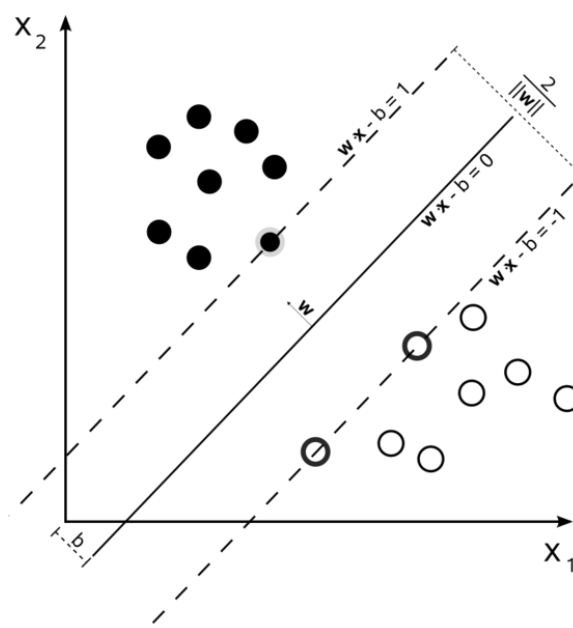
## Jak naučit počítač předvídát

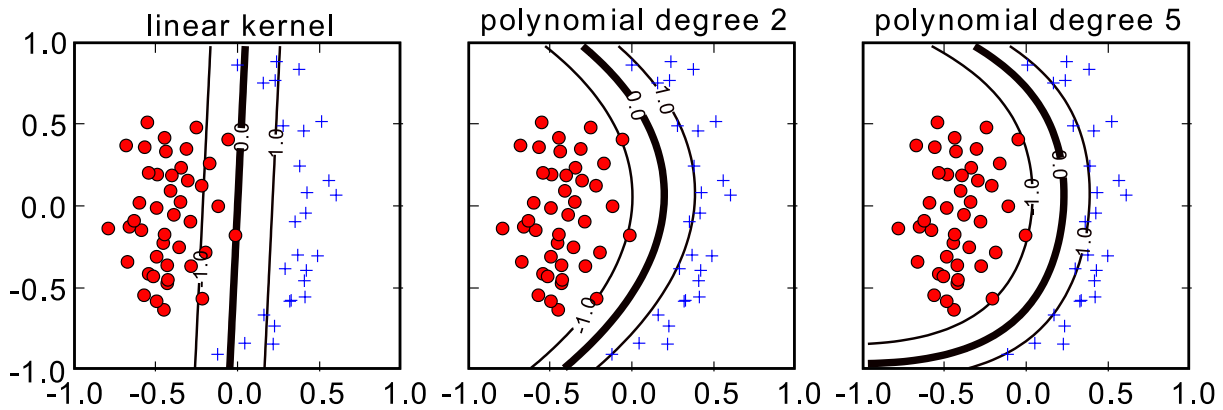
Pro začátek je nutno rozdělit data. V elementárních případech je to velmi snadné, body rozdělí přímka. Ve většině případů je situace mnohem komplikovanější. Body jsou umístěny v n-rozměrném prostoru, kde n je počet sledovaných parametrů. Body pak musíme rozdělit útvarem o (n-1) rozměrech. Proto jsou tyto výpočty prováděny v počítači.

Potřebujeme nalézt přímku, která rozděluje černé a bílé body a zároveň vzdálenost mezi přímkou a dvěma nejbližšími body (černým a bílým) je co možná největší.

U neoddělitelných dat lze modifikovat pomocí penalizační konstanty C, hodnotící překryv bodů a pomocí  $\gamma$ , které mění složitost modelu.

Na obrázku 1 jsou znázorněny body: černé značí, že jev nenastal, bílé, že jev nastal. Osy jsou určité parametry.

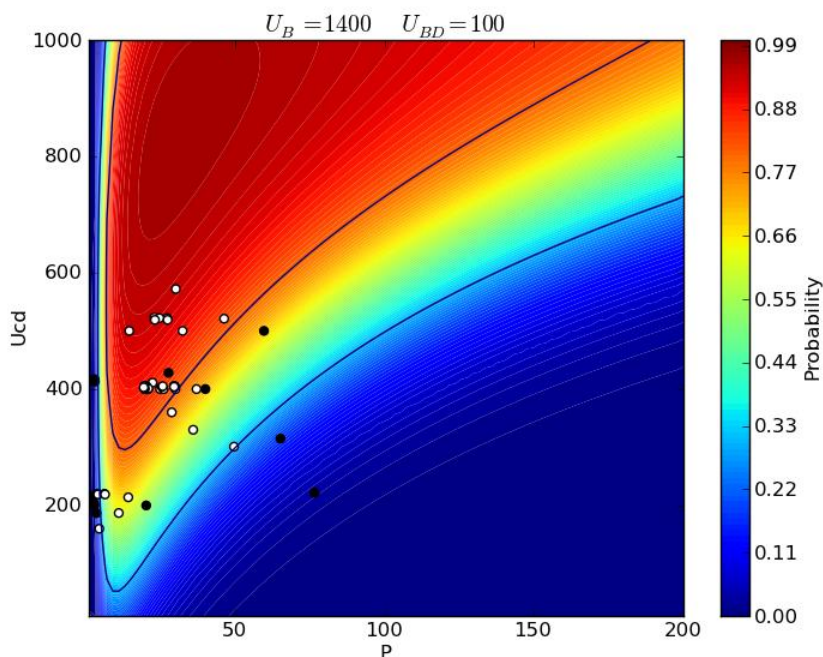




Obrázek 2: pokud body přímkou rozdělit nelze, je vhodné použít nelineární transformaci přímky.

## Pravděpodobnost vzniku plazmy v tokamaku GOLEM

V tomto konkrétním případě je rozdělení bodů velmi komplikované. GOLEM totiž neposílá parametry dva, ale devatenáct. Pro vytvoření databáze bylo zapotřebí stáhnout údaje z přibližně 2500 pokusů, vybrat pouze 6 pro nás významných parametrů, upravit je do správného formátu a vložit do programu libSVM (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>). I poté lze chování plazmatu předpovědět jen s velkými obtížemi, plazma je totiž značně nestabilní. Výsledný graf by vypadal přibližně takto.



Graf 1:

Bílé body – úspěšné výsledky

Černé body – neúspěšné výsledky

Z grafu plyne, že ač má program minimum informací o většině kombinací parametrů, je stále schopen jistého stupně predikce chování plazmatu.

## Aplikování na pokusy – výsledky

V našem miniprojektu jsme aplikovali program s databází na aktuální výsledky ostatních účastníků Týdne vědy, kteří na GOLEMovi zpracovávali jejich projekty.

Naši kolegové se ale ve svých projektech snažili dosáhnout co možná nejlepších výsledků, a

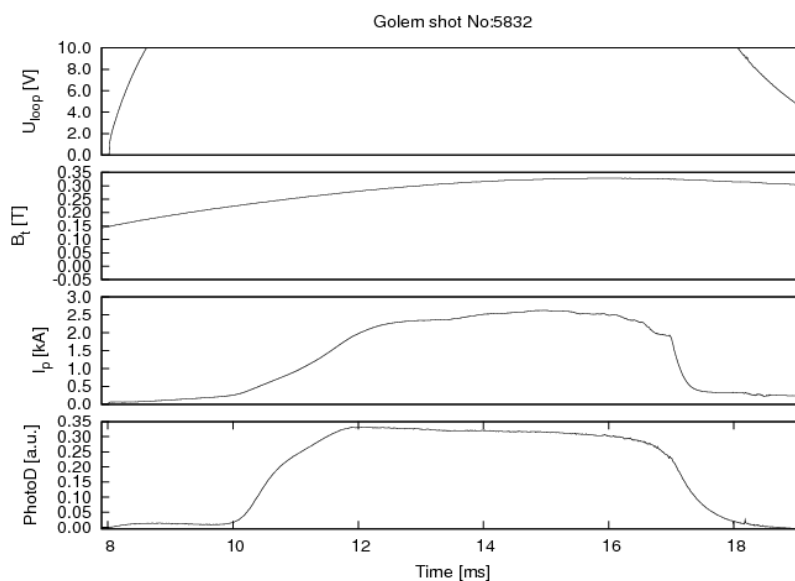
tak náš věšticí algoritmus neměl žádný problém jednoznačně určit, zda se v tokamaku objeví plazma.

**Ze 44 zkoumaných výstřelů bylo ve 100% správně odhadnuto, zda vznikne plazma.**

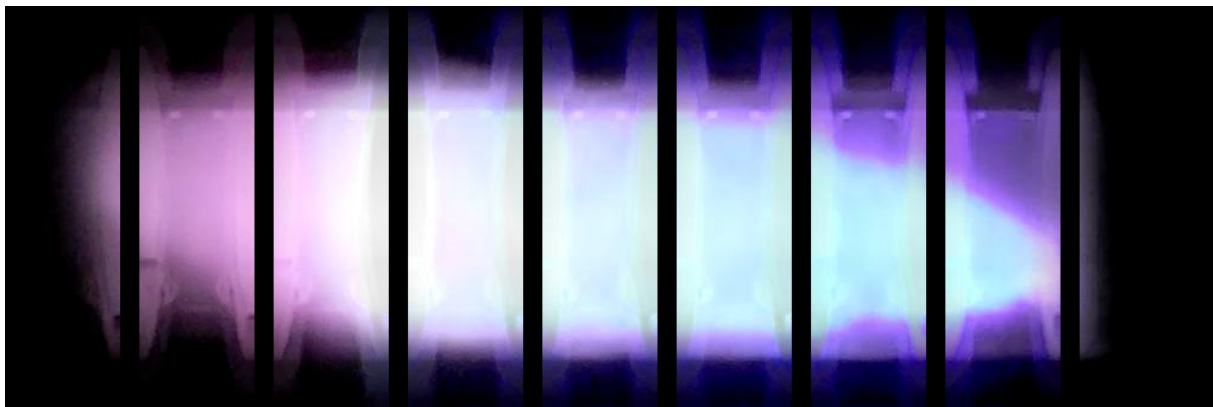
Program si byl výsledkem jistý průměrně na 93,18%.

## Pokročilé experimentování

Nakonec jsme se rozhodli vyzkoušet ty oblasti, kde podle Grafu 1 vznikne plazma s pravděpodobností 40%-60%. Během těchto pokusů jsme použili extrémní hodnoty: od velmi vysokých hodnot, které ještě nebyly na tokamaku GOLEM vyzkoušeny, až po výstřely s extrémně nízkým tlakem  $H_2$  a nízkým napětím na kondenzátorech pro vlečení proudu (current drive).



Obrázek 3: výstřel č.5832  $p_{H_2} = 100\text{mPa}$ ,  $U_{CD} = 750\text{V}$ ,  $U_{BD} = 100\text{V}$ ,  $U_B = 1400\text{V}$



Obrázek 4: fotografie výstřelu č.5832 (rychlost snímání: 1200 fps)

**Z 12 testovaných výstřelů bylo ve 100% správně odhadnuto, zda vznikne plazma.**

Program si byl výsledkem jistý průměrně na 91,7%.

## Závěr

Po zpracování výsledků tohoto projektu je možné říci, že předvídání chování tokamaku je velmi úspěšné. Skvělých výsledků se nám dostávalo i v oblastech, kde ještě žádné výstřely provedeny nebyly a program je musel odhadovat.

Tohoto výsledku bylo dosaženo i přes nedostatek dat v databázi programu. Pokud bychom měli k dispozici více dat, dosáhli bychom větší jistoty předpovědi.

## Poděkování

V první řadě bychom rádi poděkovali našemu supervizorovi, Michalovi Odstrčilovi, který nás naučil základy práce s Linuxy, základy programování v jazyku Python, tvorbu grafů v programu Gnuplot, práci v grafickém editoru Inkscape a věnoval nám spoustu času.

Dále bychom rádi poděkovali Ing. Vojtěchu Svobodovi, CSc., za jeho vstřícnost k naším extrémně destruktivním požadavkům, všem organizátorům TV@J, FJFI na ČVUT za propůjčení zařízení a zázemí k experimentům.

Nakonec bychom rádi poděkovali všem známým i neznámým sponzorům akce TV@J.

## Reference:

- [1] Ben-Hur, A. - Weston, J.: *A User's Guide to Support Vector Machines* 2011,
- [2] Chih-Wei Hsu - Chih-Chung Chang - Chih-Jen Lin: *A Practical Guide to Support Vector Classification* 2010
- <http://golem.fjfi.cvut.cz/>