Alex Smola

Barnabas Poczos

Ina Fiterau

CMU - MLD

Examples

LIBSVM

LIBSVM

http://www.csie.ntu.edu.tw/~cjlin/libsvm/
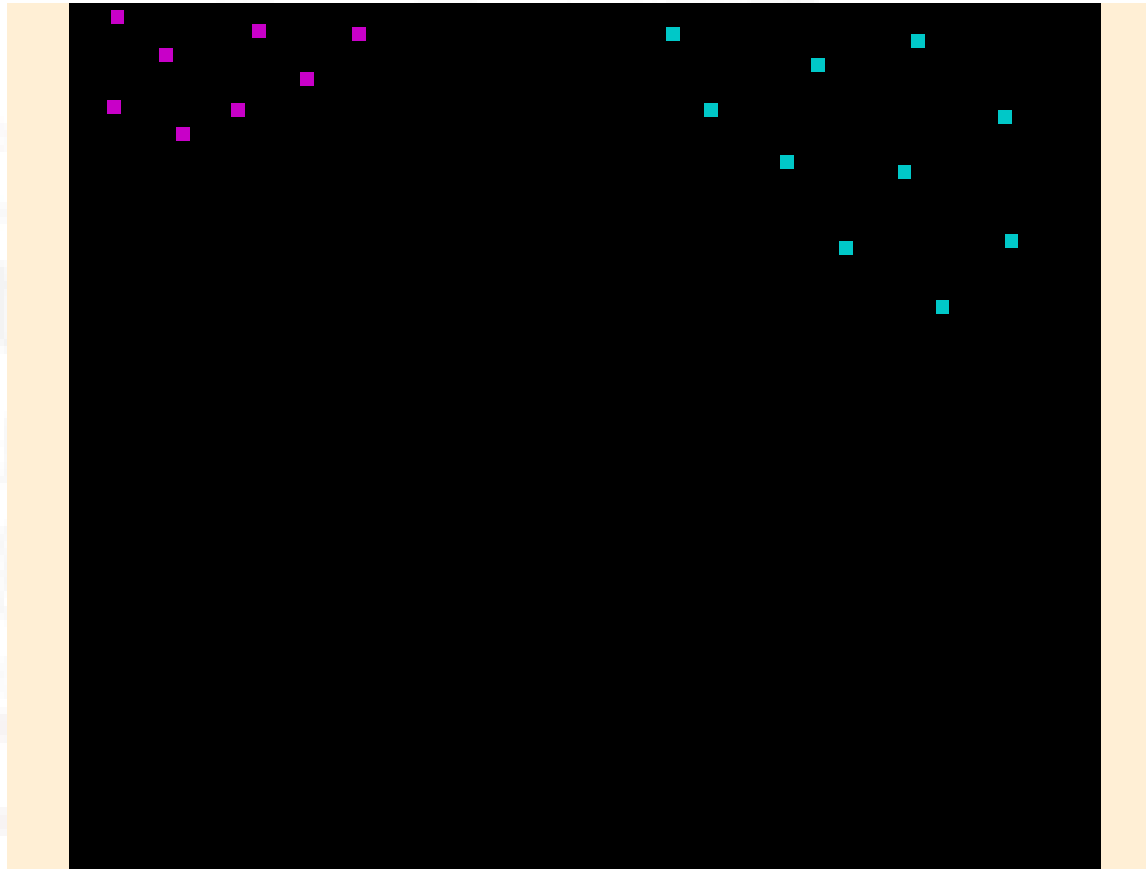
- Multiple interfaces and extensions
  - Java, Matlab, Python, R, C#
- DO NOT use in the homework, though.
  - We want you to learn what's under the hood & how you can tweak it to make it awesome
- DO play with the applet ☺
- If you're running Debian:

```
> apt-get install libsvm
```

Happy 2-class Problem

Q: Just how easy is this?

Happy 2-class Problem

Q: Just how easy is this?

A: Piece of cake!

Happy 2-class Problem

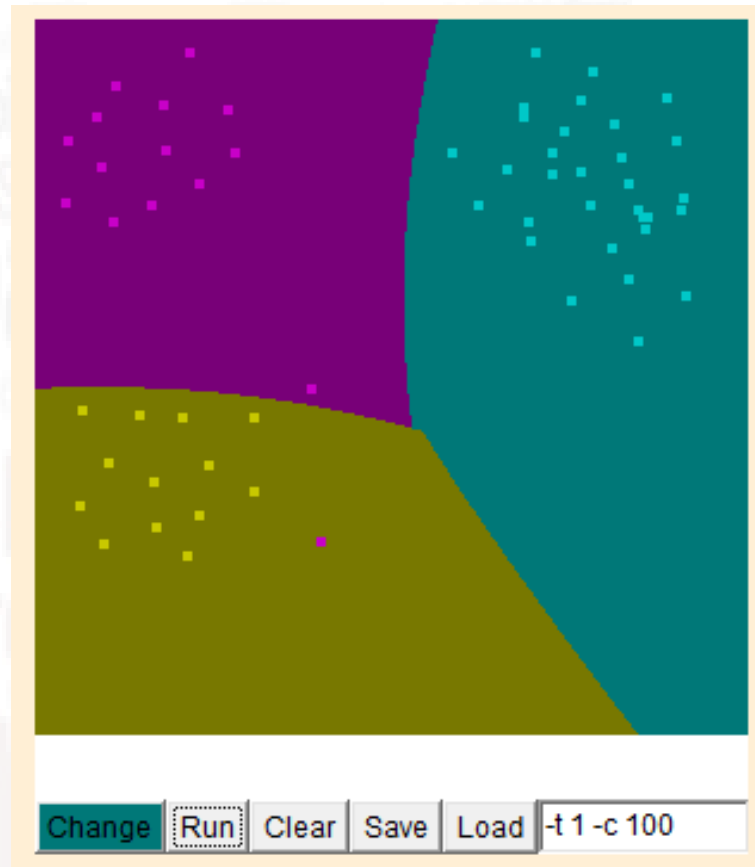Q: What happens when I add points?

Happy 2-class Problem

Q: What happens when I add points?

A: Within the cvx hull of existing ones (same class): NOTHING

Happy 2-class Problem

Q: What happens when I add points?

A: Outside of cvx hull, boundary changes to maintain margin.

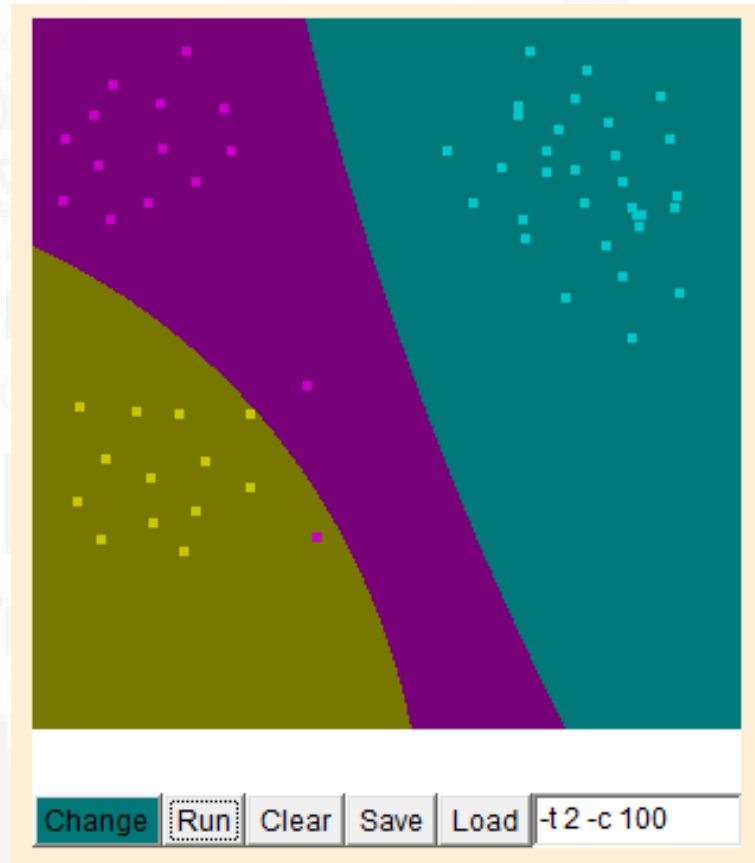| Change | Run | Clear | Save | Load | -t 0 -c 100 |

Still Happy 3-class Problem

Q: What happens when we change kernels?

Still Happy 3-class Problem

Q: What happens when we change kernels?

A: This. [Polinomial]

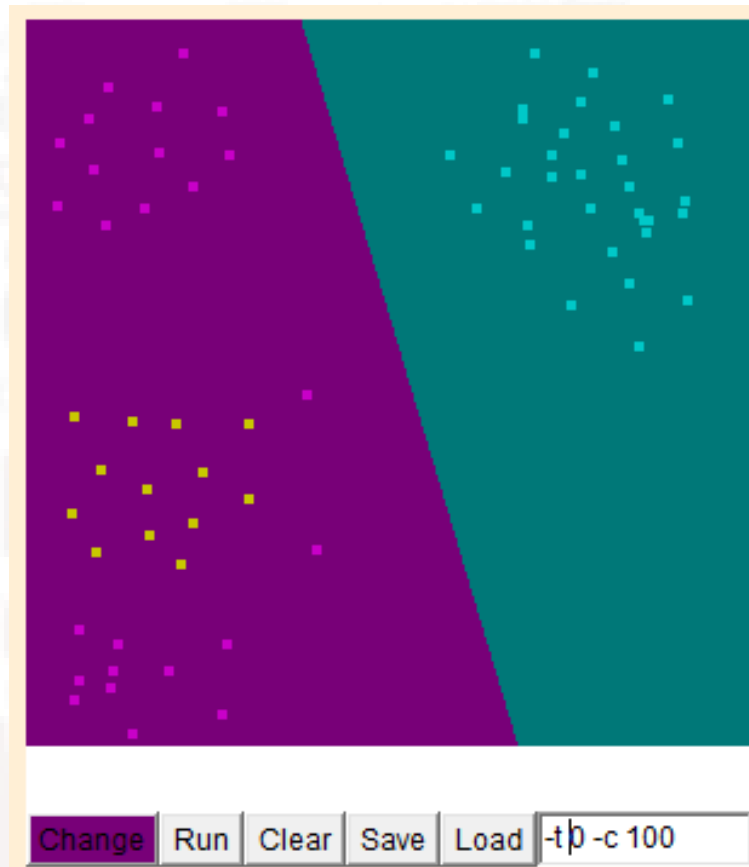## Still Happy 3-class Problem

Q: What happens when we change kernels?
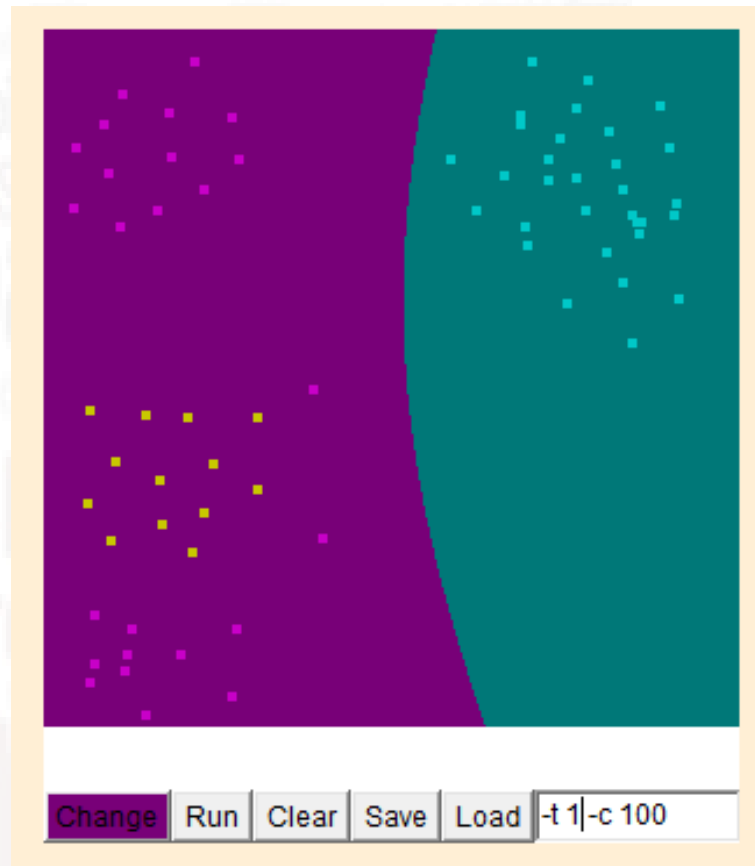
A: This. [Gaussian]

Still Happy 3-class Problem

Q: What happens when we change kernels?

A: This. [Sigmoid]
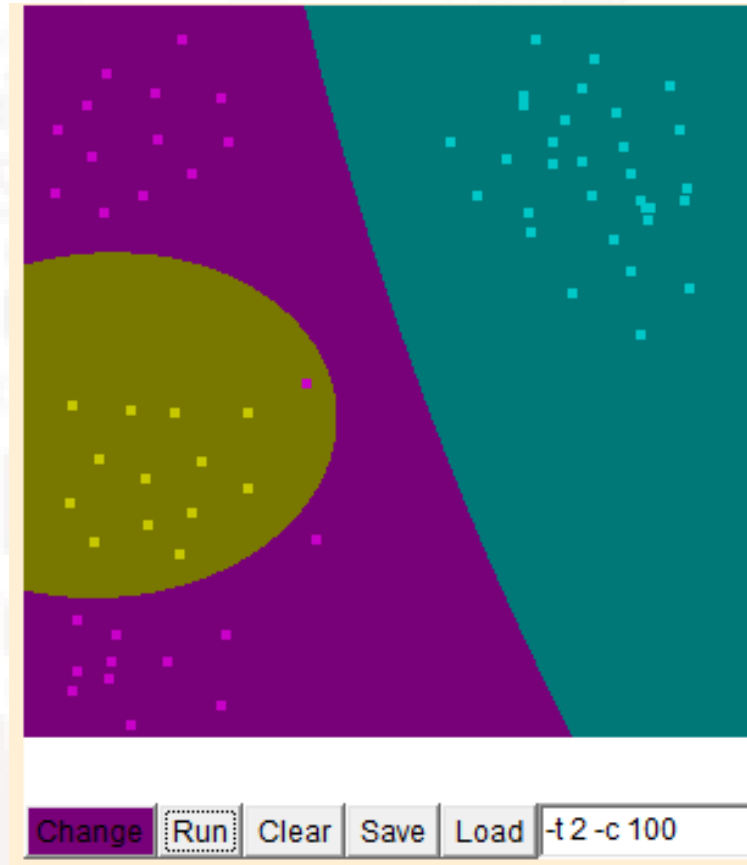
3-Class Problem

Q: This is what a linear kernel does. What happens if I change it?

3-Class Problem

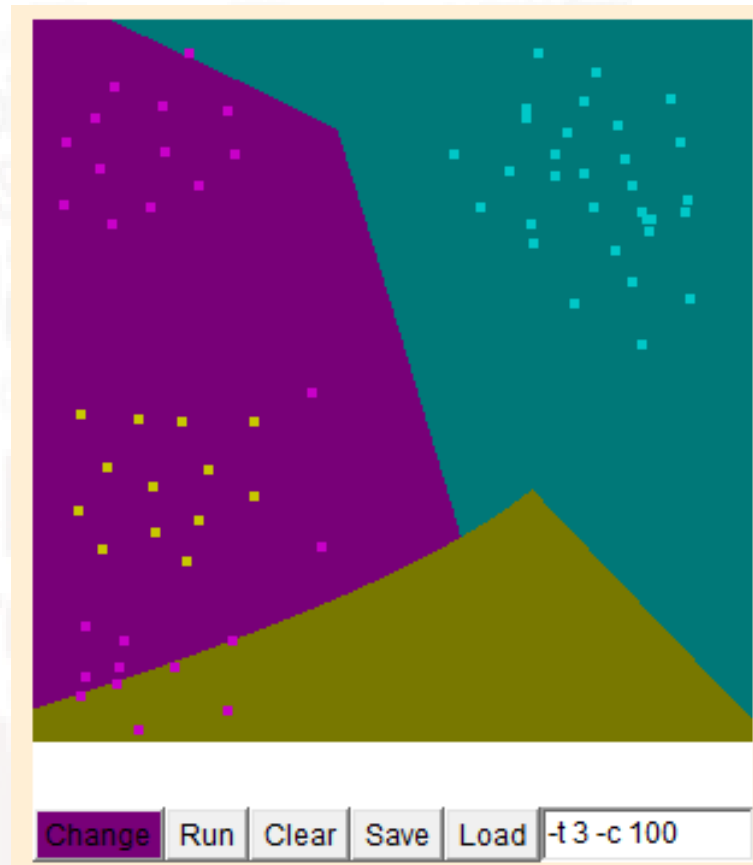Q: This is what a linear kernel does. What happens if I change it?

A: Polynomial. Doesn't solve the problem.

Change | Run | Clear | Save | Load | -t 2 -c 100

3-Class Problem
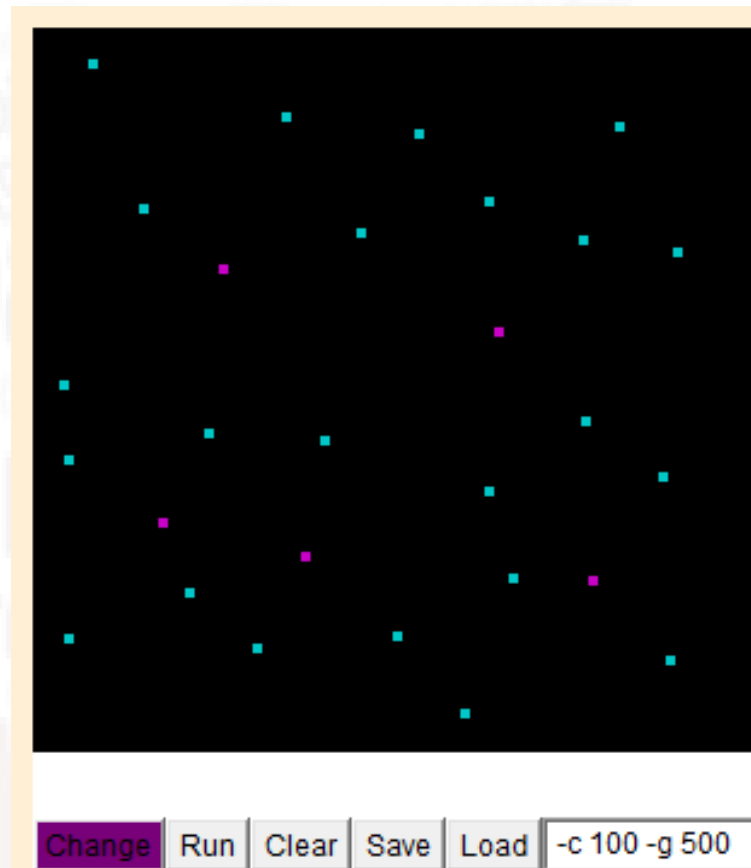
Q: This is what a linear kernel does. What happens if I change it?

A: Gaussian. OK!

3-Class Problem

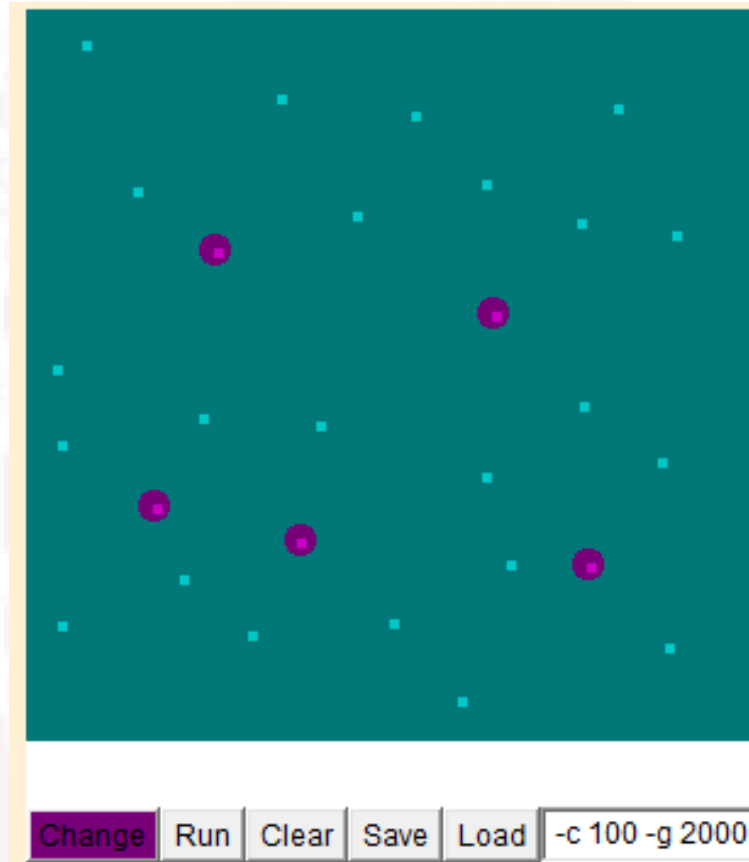Q: This is what a linear kernel does. What happens if I change it?

A: Sigmoid. (Weird, huh?)

| Change | Run | Clear | Save | Load | -c 100 -g 500 |

Gaussian Kernel

Q: What makes it different?

Change | Run | Clear | Save | Load | -c 100 -g 2000

Gaussian Kernel

Q: What makes it different?

A: The capability of its basis function to 'focus' on one point.

LIBSVM Parameters: -t 2 -c 100 -g 2000

Gaussian Kernel

Q: What makes it different?

A: The capability of its basis function to 'focus' on one point.

LIBSVM Parameters: -t 2 -c 100 -g 500

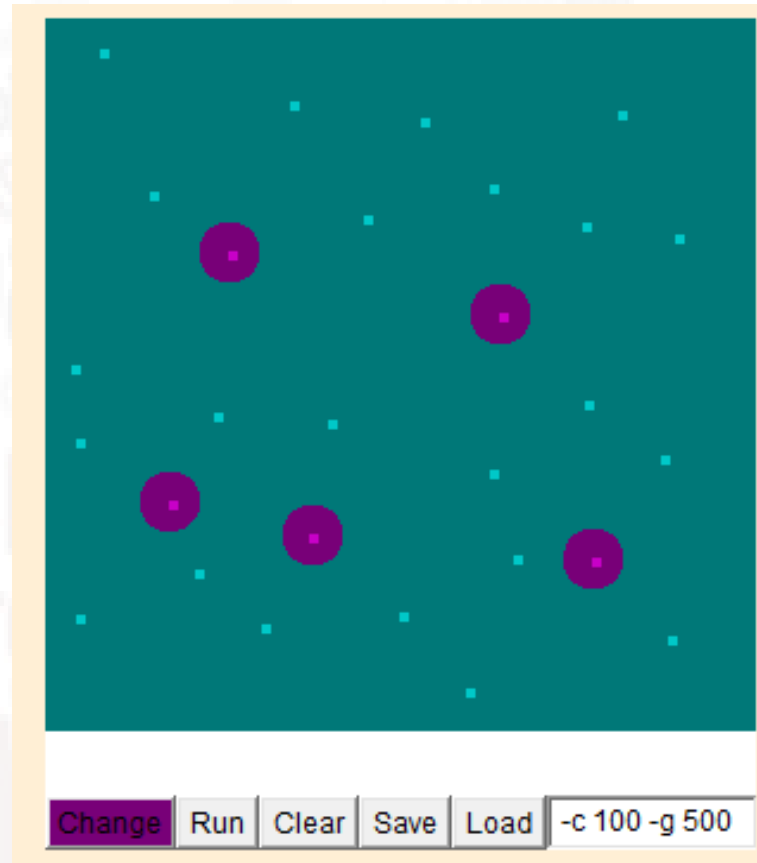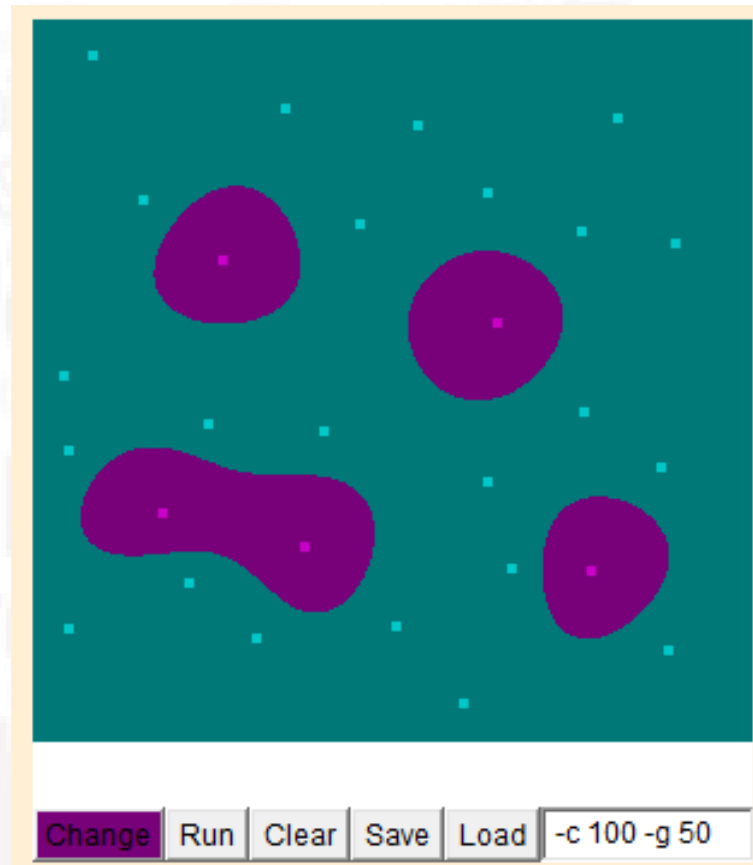| Change | Run | Clear | Save | Load | -c 100 -g 50 |

Gaussian Kernel

Q: What makes it different?

A: The capability of its basis function to 'focus' on one point.

LIBSVM Parameters: -t 2 -c 100 -g 50

Change | Run | Clear | Save | Load | -t 2 -c 100 -g 7

Gaussian Kernel

Q: What makes it different?

A: The capability of its basis function to 'focus' on one point.

LIBSVM Parameters: -t 2 -c 100 –g 7

| Change | Run | Clear | Save | Load | -t 2 -c 100 -g 5 |

Gaussian Kernel

Q: What makes it different?

A: The capability of its basis function to 'focus' on one point.
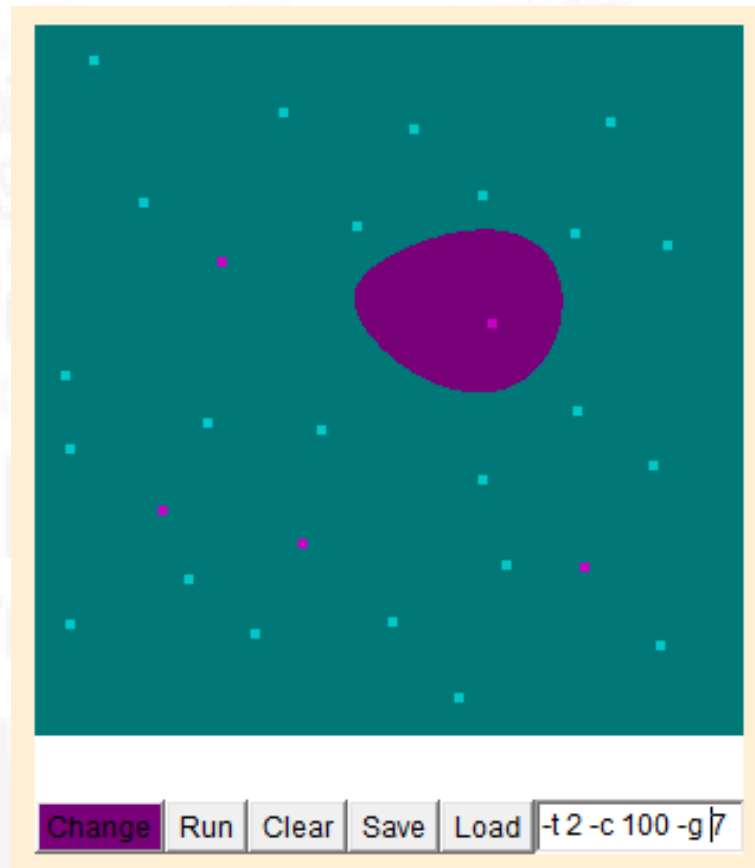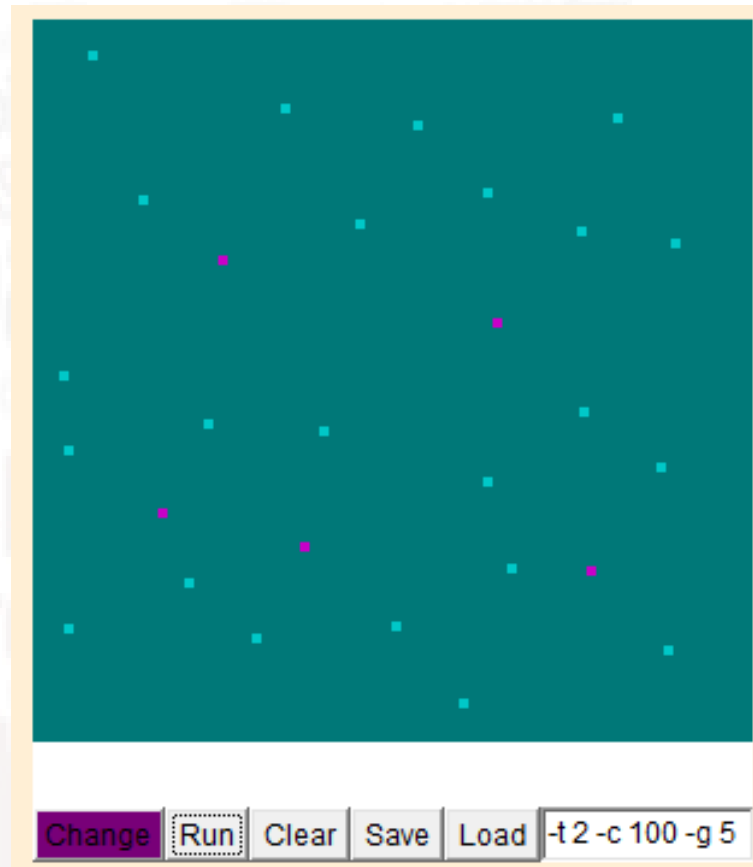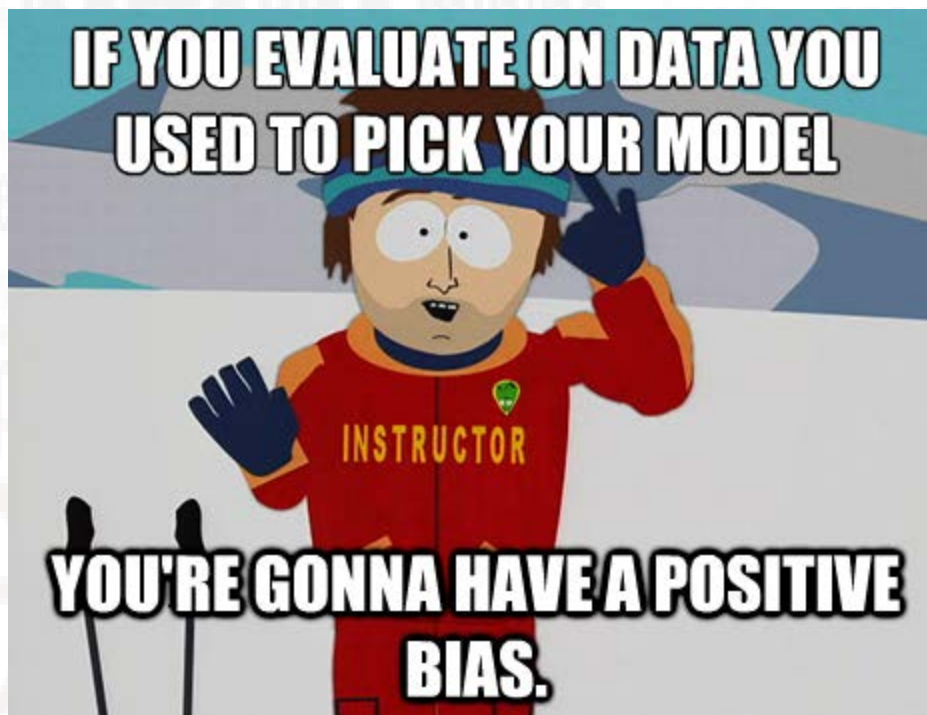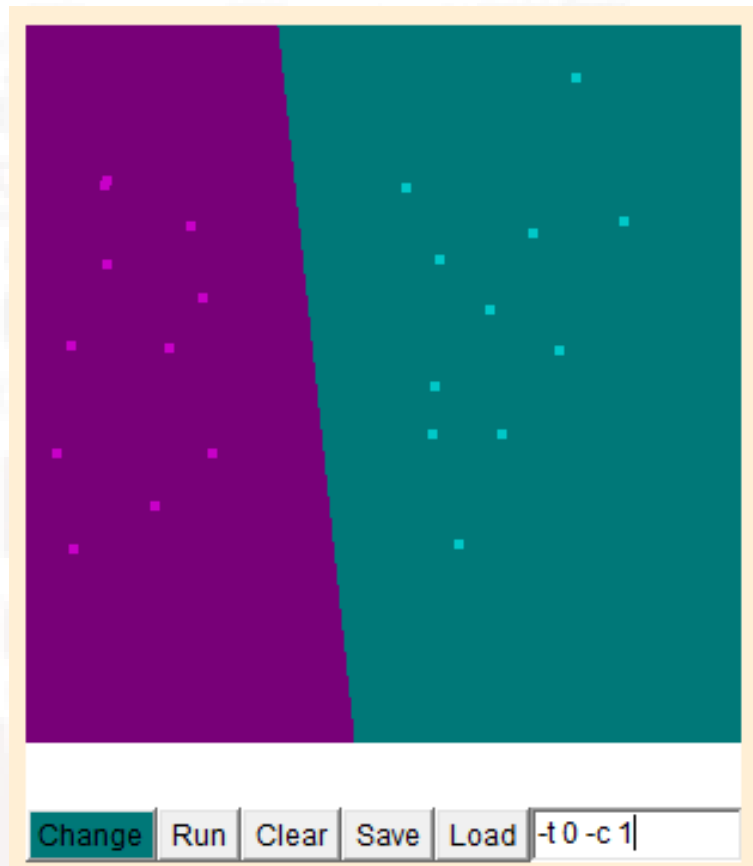
LIBSVM Parameters: -t 2 -c 100 -g 5

Get correct parameter through cross-validation!

Careful how you cross-validate

| Change | Run | Clear | Save | Load | -t 0 -c 1 |

Happy 2-class problem

Q: How does the penalty C influence the border?

A: It doesn't.

| Change | Run | Clear | Save | Load | -t 0 -c 1000 |

Happy 2-class problem

Q: How does the penalty C influence the border?

Unhappy 2-class problem

Q: How does the penalty C influence the border?

A: For high C, you get low tolerance for outliers (high sensitivity)

`Change | Run | Clear | Save | Load | -t 0 -c 1`

Unhappy 2-class problem

Q: How does the penalty C influence the border?

A: For low C, you get high tolerance for outliers (high robustness)

Conclusions

- Convex hulls dictate borders for non-overlapping classes
- Type of kernel has HUGE impact
  - Pick depending on application.
- Kernel parameters matter
  - Get correct value through cross-validation
- Slack penalty influences robustness