Fruit Fly Optimization

Presenter: Hasan Masroor

Fall Semester 2025

“Evolving support vector machines using fruit fly optimization for medical data classification”

Liming Shen, Huiling Chena, Zhe Yu, Wenchang Kang, Bingyu Zhang, Huaizhong Li, Bo Yang, DayouLiu

Knowledge-Based Systems (Elsevier)

doi.org/10.1016/j.knosys.2016.01.002

Year: 2016

Keywords:

Support vector machine
Parameter optimization
Fruit fly optimization
Medical diagnosis

Knowledge-Based Systems 96 (2016) 61–75

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

ELSEVIER

Evolving support vector machines using fruit fly optimization for medical data classification

CrossMark

Liming Shen[a], Huiling Chen[a,c,*], Zhe Yu[a], Wenchang Kang[a], Bingyu Zhang[a], Huaizhong Li[a], Bo Yang[b,c], Dayou Liu[b,c]

[a] *College of Physics and Electronic Information Engineering, Wenzhou University, 325035 Wenzhou, China*
[b] *College of Computer Science and Technology, Jilin University, Changchun 130012, China*
[c] *Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China*

ARTICLE INFO

Article history:
Received 7 March 2015
Revised 31 December 2015
Accepted 2 January 2016
Available online 11 January 2016

Keywords:
Support vector machine
Parameter optimization
Fruit fly optimization
Medical diagnosis

ABSTRACT

In this paper, a new support vector machines (SVM) parameter tuning scheme that uses the fruit fly optimization algorithm (FOA) is proposed. Termed as FOA-SVM, the scheme is successfully applied to medical diagnosis. In the proposed FOA-SVM, the FOA technique effectively and efficiently addresses the parameter set in SVM. Additionally, the effectiveness and efficiency of FOA-SVM is rigorously evaluated against four well-known medical datasets, including the Wisconsin breast cancer dataset, the Pima Indians diabetes dataset, the Parkinson dataset, and the thyroid disease dataset, in terms of classification accuracy, sensitivity, specificity, AUC (the area under the receiver operating characteristic (ROC) curve) criterion, and processing time. Four competitive counterparts are employed for comparison purposes, including the particle swarm optimization algorithm-based SVM (PSO-SVM), genetic algorithm-based SVM (GA-SVM), bacterial forging optimization-based SVM (BFO-SVM), and grid search technique-based SVM (Grid-SVM). The empirical results demonstrate that the proposed FOA-SVM method can obtain much more appropriate model parameters as well as significantly reduce the computational time, which generates a high classification accuracy. Promisingly, the proposed method can be regarded as a useful clinical tool for medical decision making.

- SVM performance depends heavily on parameter tuning (C, γ).
- Traditional tuning (grid-search, gradient-descent) is slow and prone to local minima.
- Metaheuristics (GA, PSO, BFO) are effective but complex and costly.
- FOA: a simple and fast nature-inspired optimizer.
- Goal: use FOA to optimize SVM (FOA-SVM) for medical datasets.

➢ FOA: Applications and Improved Versions

- Widely used for parameter optimization in ML models

(SVM, forecasting models, etc)

- Applied in real-world prediction and engineering problems

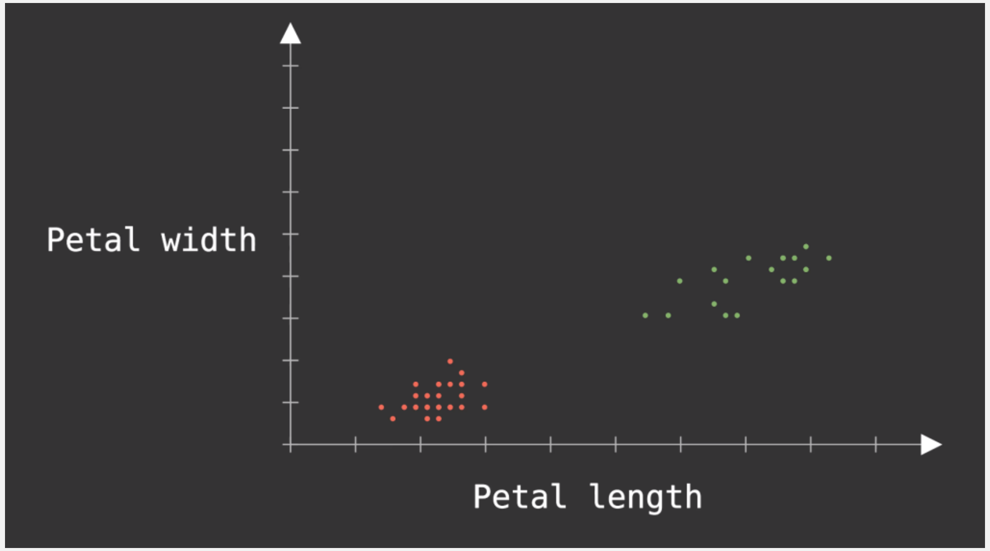
(electric load forecasting, satisfaction detection, power load modeling, steel casting)

- Many enhanced FOA variants developed

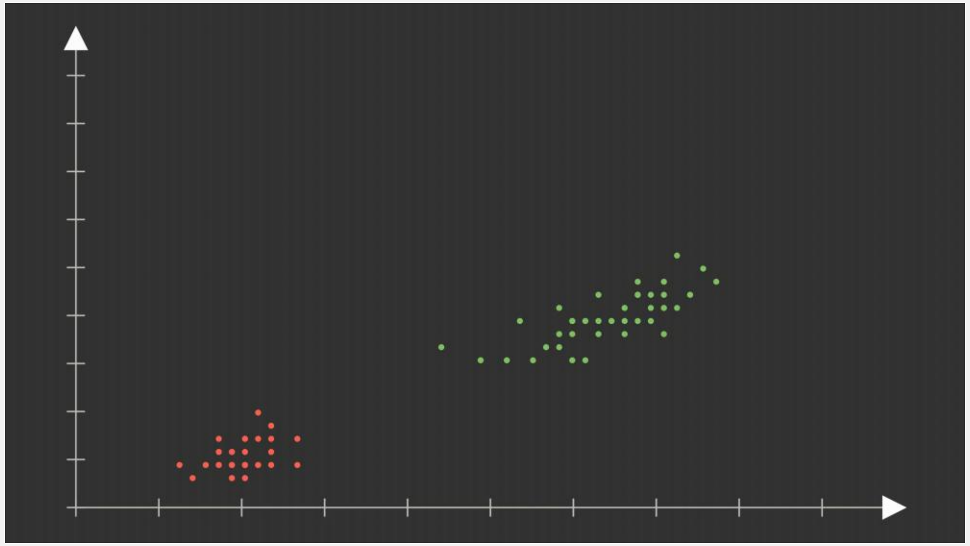
(bFOA, MFOA, IFOA, AM-FOA, CFOA, nFOA, etc)

دانشگاه صنعتی امیرکبیر

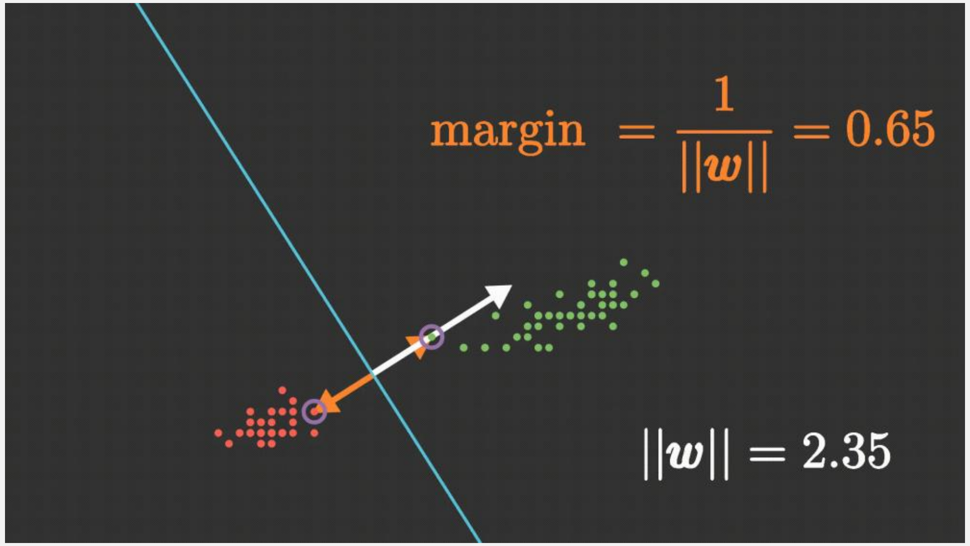
1



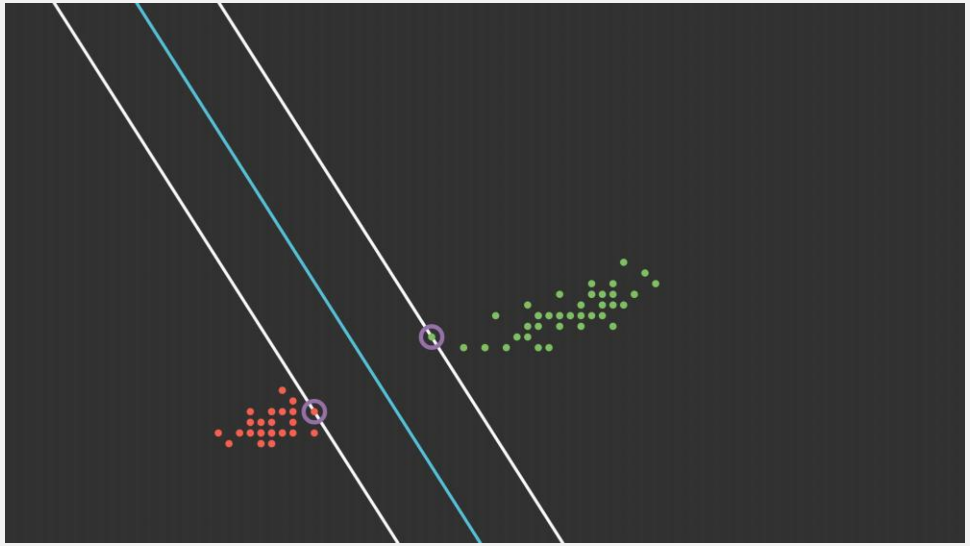
2



3



4



Support vector machines (SVM)

h:X → {−1,+1}

$$\begin{cases} \boldsymbol{w}.\boldsymbol{x}_+ + b \geq 1 & (y=+1) \\ \boldsymbol{w}.\boldsymbol{x}_- + b \leq -1 & (y=-1) \end{cases}$$

$$\Rightarrow y_i\ (\boldsymbol{w}.\boldsymbol{x_i} + b) - 1 \geq 0, \forall i$$

(1)

$$\mathrm{Min}g(\mathrm{w},\xi) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i$$

$$s.t.,\ y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, \xi_i \geq 0$$

(2)

$$\max_\alpha \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j \mathbf{x}_i^T\mathbf{x}_j$$

$$s.t.,\ \alpha_i \geq 0, i = 1,\ldots,n,\ \sum_{i=1}^{n}\alpha_i y_i = 0$$

(3)

$$g(\mathbf{x}) = \mathrm{sgn}\left(\sum_{i=1}^{n}\alpha_i y_i \mathbf{x}_i^T\mathbf{x} + b\right)$$

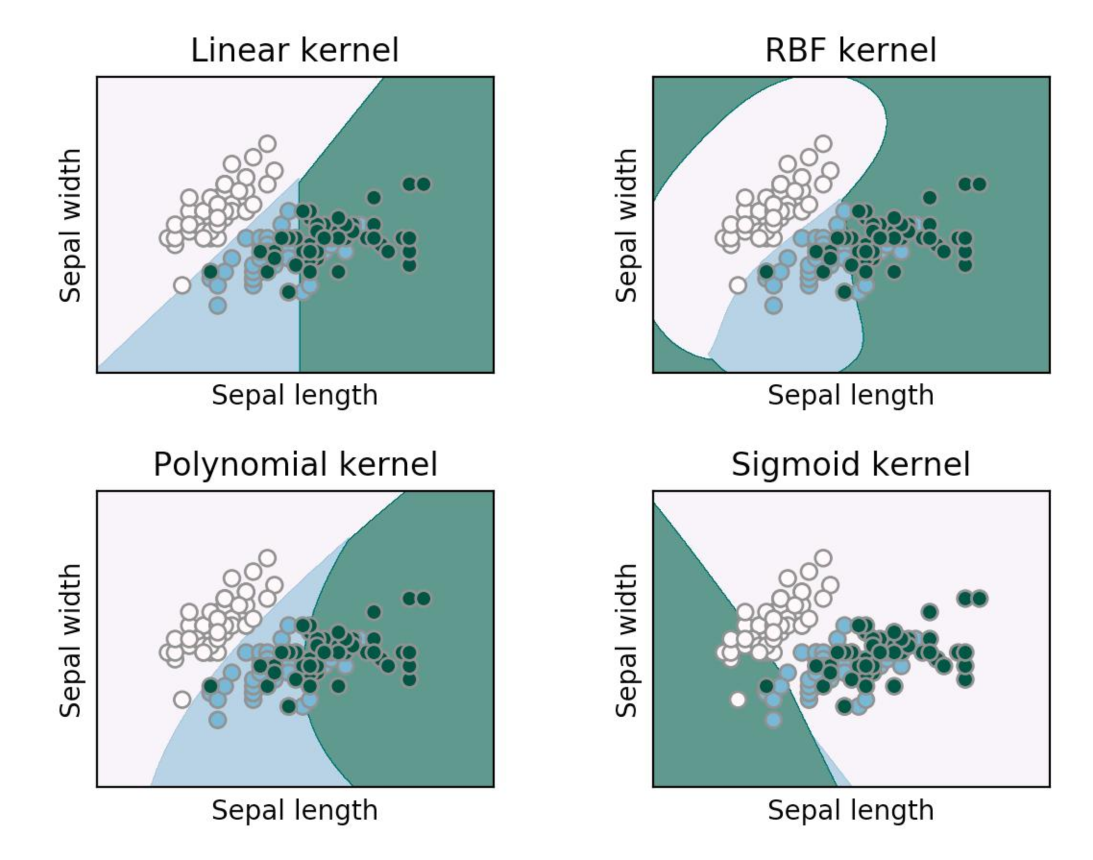
(4)

➢ Support vector machines (SVM)

$$g(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i \phi(\mathbf{x}_i)^T \phi(\mathbf{x}) + b\right) \quad (5)$$

$$g(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b\right) \quad (6)$$

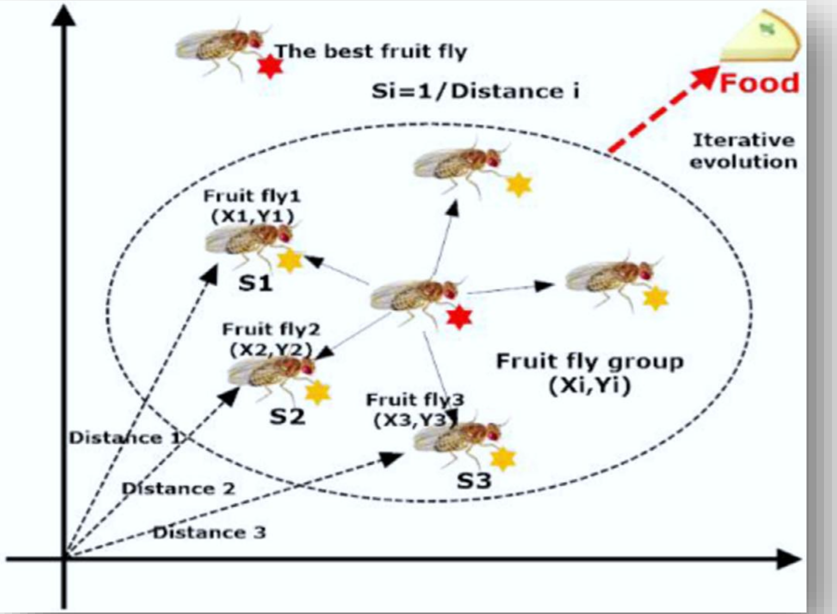
Kernel types	Kernel functions
Linear kernel	$K(x, x_i) = (x^T x_i)$
Polynomial kernel	$K(x, x_i) = ((x^T x_i) + 1)^d$
Radial based kernel (RBF)	$K(x, x_i) = \exp(-\gamma \lVert x - x_i \rVert^2)$
Sigmoid kernel	$K(x, x_i) = \tanh((x^T x_i) + b)$



- **What FOA Is Inspired By**

  Inspired by fruit-fly foraging behavior

- **Core Idea**



- **Why FOA?**

i. Fast convergence

ii. Easy to implement compared to GA/PSO/BFO

iii. Very simple, few parameters

iv. Effective for parameter optimization tasks

- **Step 1:** Parameters initialization

$$X\_axis = \mathrm{rands}(1,\ 2)$$

$$Y\_axis = \mathrm{rands}(1,\ 2)$$

- **Step 2:** Population initialization

$$X_i = X\_axis + RandomValue$$

$$Y_i = Y\_axis + RandomValue$$

- **Step 3:** Population evaluation

$$D_i = \sqrt{X_i^2 + Y_i^2}$$

$$S_i = 1/D_i$$

- **Step 4:** Replacement

$$Smell_i = \text{Function}(S_i)$$

- **Step 5:** Find the maximal smell concentration

$$[bestSmell bestIndex] = max(Smell)$$

- **Step 6:** Keep the maximal smell concentration

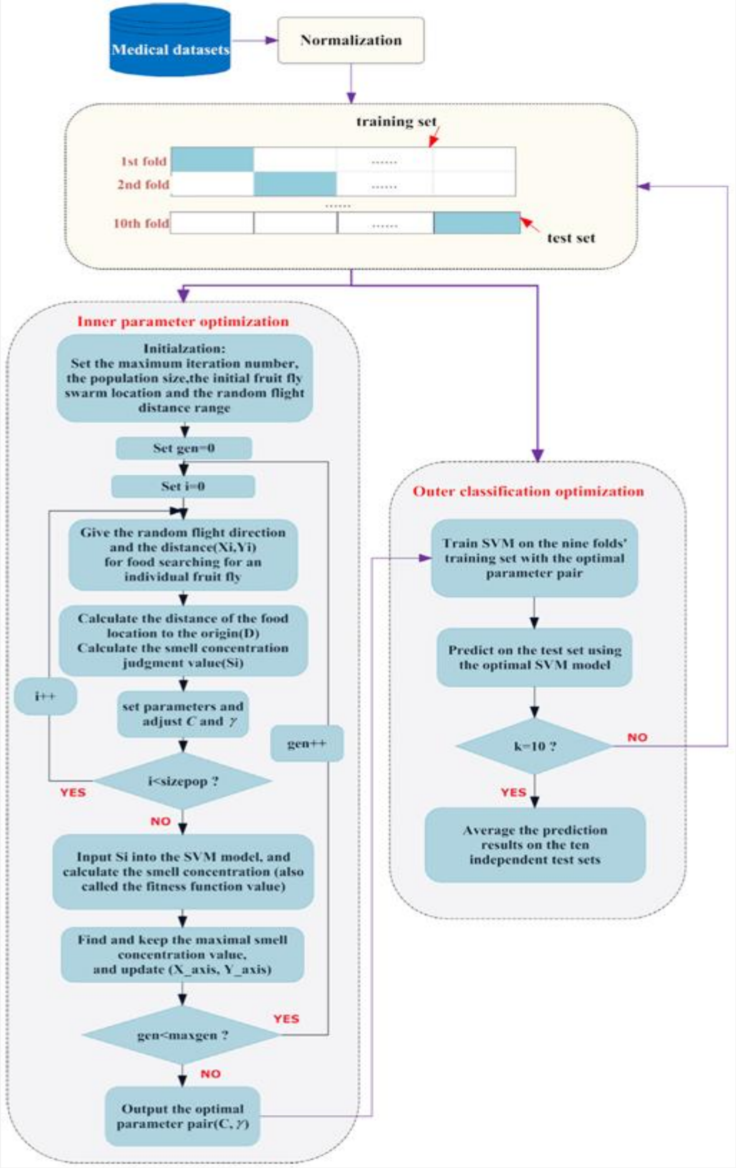
$$Smellbest = bestSmell$$

$$X\_axis = X(bestIndex)$$

$$Y\_axis = Y(bestIndex)$$

- **Step 7:** Iterative optimization

- **Flowchart of the proposed FOA-SVM method**



Datasets & evaluation metrics

- Accuracy = TP+TN/(TP+FP+FN+TN) × 100%
- Sensitivity = TP/(TP+FN) × 100%
- Specificity = TN/(FP+TN) × 100%
- AUC

No.	Datasets	# of classes	# of instances	# of features	Miss.
1	Wisconsin breast cancer (Wisconsin)	2	699	9	Yes
2	Pima Indians diabetes (Pima)	2	768	8	No
3	Parkinson	2	195	22	No
4	Thyroid	3	215	5	No

1. Breast cancer diagnosis problem

Attribute	Description	Domain
F_1	Clump thickness	1–10
F_2	Uniformity of cell size	1–10
F_3	Uniformity of cell shape	1–10
F_4	Marginal adhesion	1–10
F_5	Single epithelial cell size	1–10
F_6	Bare nuclei	1–10
F_7	Bland chromatin	1–10
F_8	Normal nucleoli	1–10
F_9	Mitoses	1–10

- 699 instances and 9 attributes
- The goal is to discriminate between the benign and malignant samples.
- Statistical test (Paired t-test):
  - ☑ $p < 0.05$
  - ☒ $p \geq 0.05$

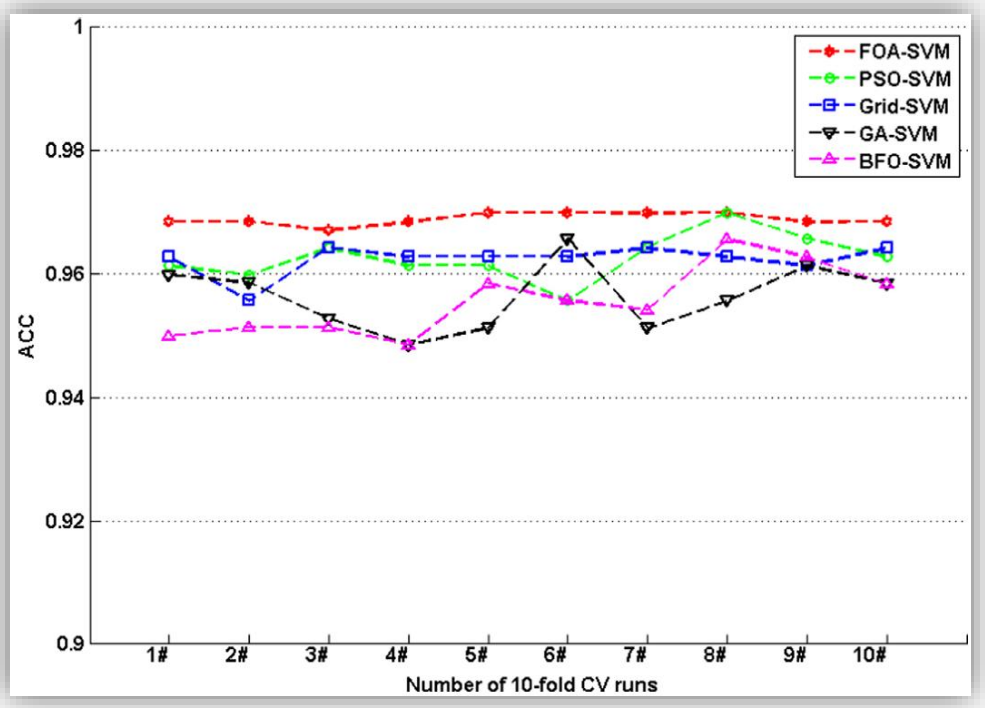
Metrics	*t*-value (significance)			
	PSO-SVM	Grid-SVM	GA-SVM	BFO-SVM
ACC	**5.038(0.001)**	**7.916(0.000)**	**7.416(0.000)**	**7.990(0.000)**
AUC	**4.486(0.002)**	**7.617(0.000)**	**8.078(0.000)**	**5.717(0.000)**
Sensitivity	**2.732(0.023)**	**3.736(0.005)**	**3.664(0.005)**	**9.993(0.000)**
Specificity	1.430(0.187)	**6.786(0.000)**	**8.228(0.000)**	**−3.469(0.007)**

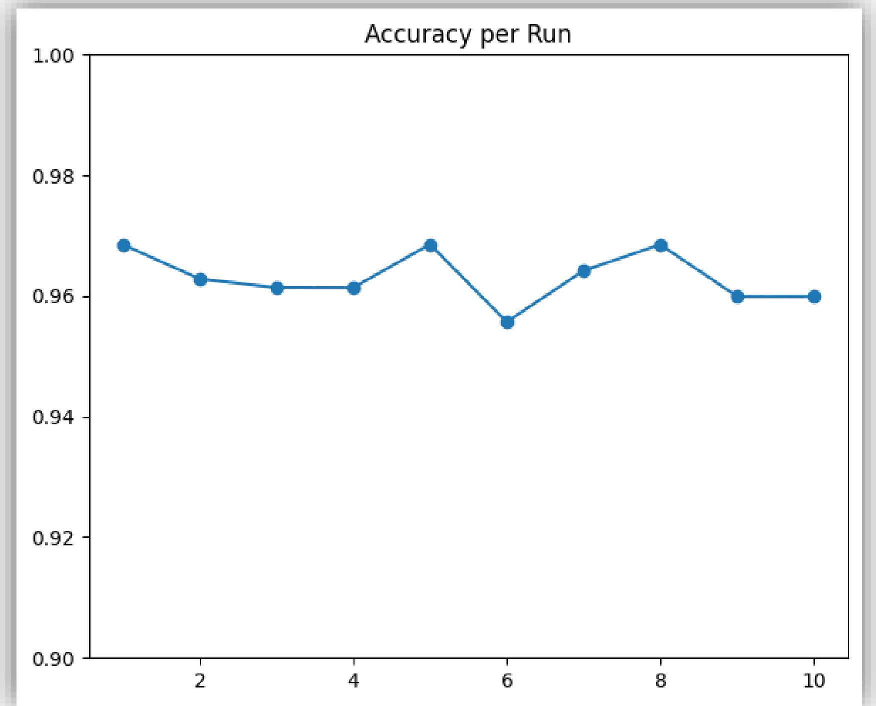
- **Performance Comparison on Wisconsin Dataset**

Metrics	Methods				
	PSO-SVM	Grid-SVM	GA-SVM	BFO-SVM	FOA-SVM
ACC	0.9627 ± 0.0038	0.9624 ± 0.0025	0.9564 ± 0.0054	0.9557 ± 0.0056	0.9690 ± 0.0010
AUC	0.9626 ± 0.0042	0.9602 ± 0.0032	0.9529 ± 0.0063	0.9609 ± 0.0044	0.9687 ± 0.0009
Sensitivity	0.9624 ± 0.0070	0.9662 ± 0.0024	0.9627 ± 0.0050	0.9457 ± 0.0082	0.9686 ± 0.0014
Specificity	0.9659 ± 0.0068	0.9545 ± 0.0060	0.9432 ± 0.0099	0.9761 ± 0.0059	0.9689 ± 0.0018

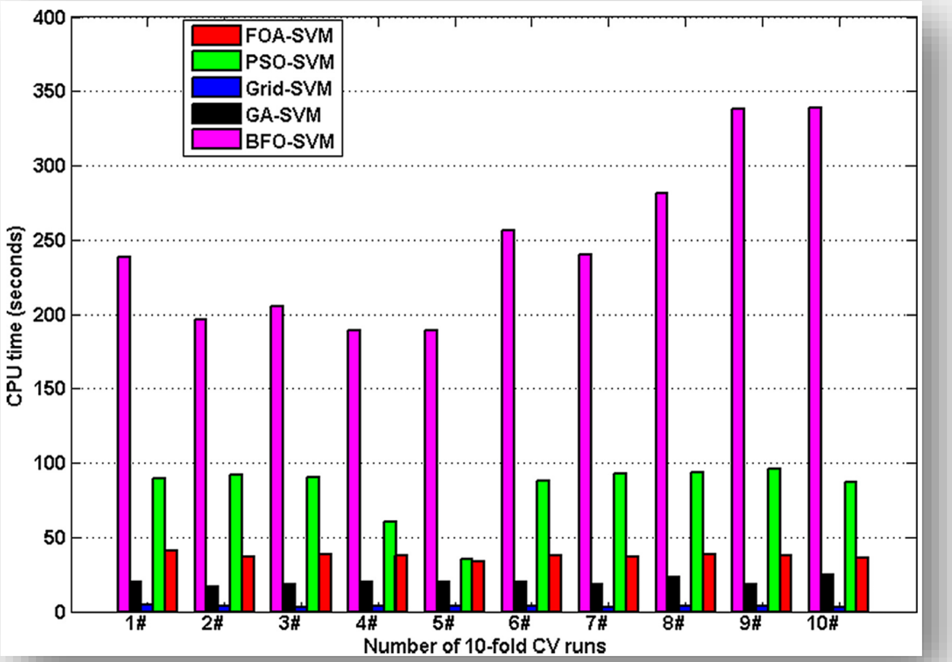
```
Accuracy: 0.9631 ± 0.0224
AUC: 0.9939 ± 0.0078
Sens: 0.9461 ± 0.0461
Spec: 0.9721 ± 0.0284
```

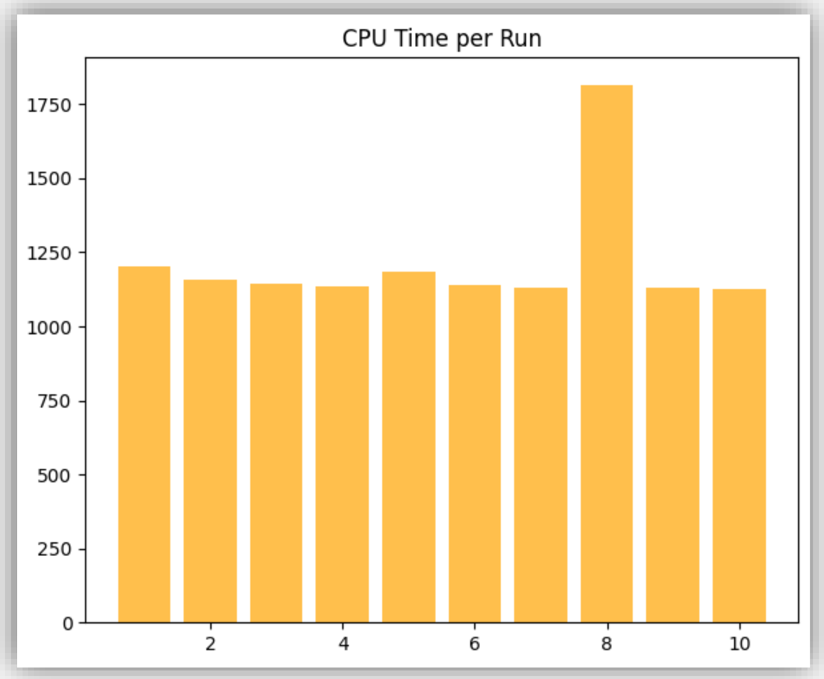
- ## Accuracy Across 10 Runs





- CPU Time Comparison – Wisconsin Breast Cancer Dataset





2. Diabetes disease diagnosis problem

- 768 instances (500 normal, 268 diabetes) and 8 attributes
- FOA-SVM achieves the best overall performance
- Lowest computation time among metaheuristic methods

Attribute no.	Attribute
1	Number or times pregnant (NTP)
2	Plasma glucose concentration (PGC)
3	Diastolic blood pressure (mmHg) (DBP)
4	Triceps skin-fold thickness (mm) (TSFT)
5	2-h serum insulin (mu U/mL) (H2SI)
6	Body mass index (kg/m2) (BMI)
7	Diabetes pedigree function (DPF)
8	Age

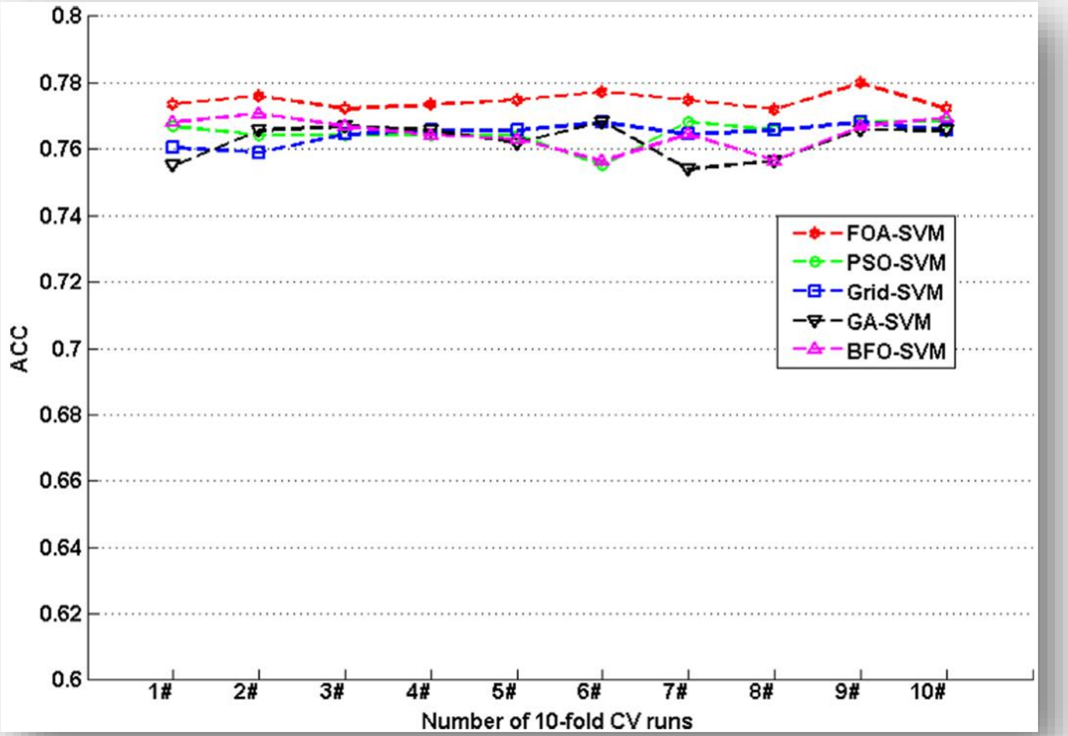
Metrics	*t*-value (significance)			
	PSO-SVM	Grid-SVM	GA-SVM	BFO-SVM
ACC	**5.985(0.000)**	**9.425(0.000)**	**7.372(0.000)**	**5.596(0.000)**
AUC	**5.566(0.000)**	**5.928(0.000)**	**6.898(0.000)**	**6.167(0.000)**
Sensitivity	1.965(0.081)	**2.707(0.024)**	**2.934(0.017)**	**3.476(0.007)**
Specificity	**3.971(0.003)**	**3.580(0.006)**	**5.690(0.000)**	**3.285(0.009)**

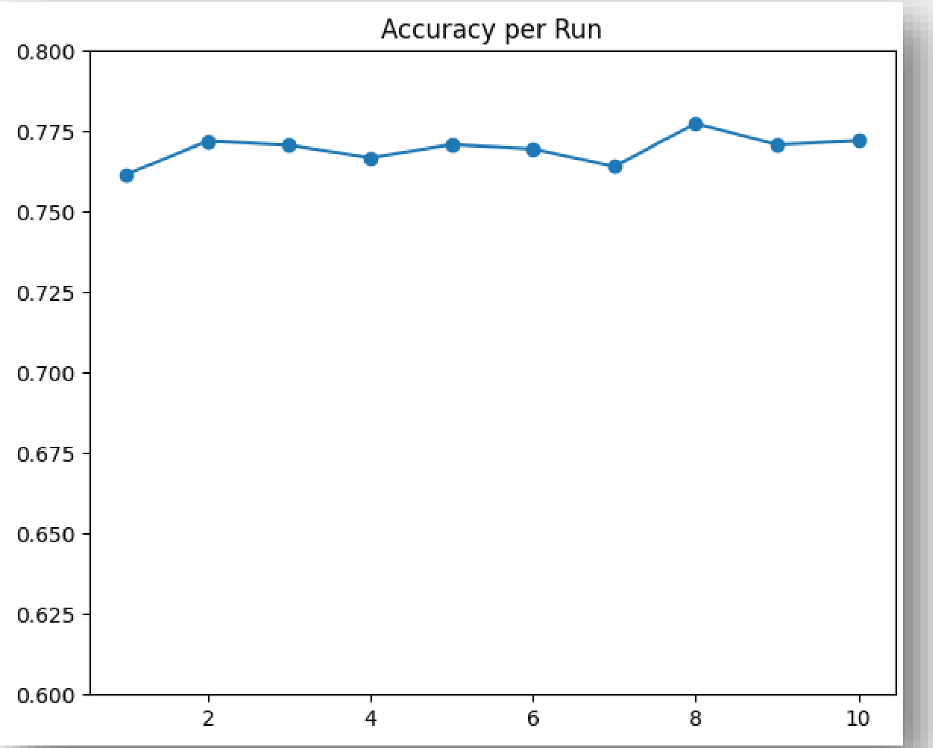
- **Performance Comparison on Pima Dataset**

Metrics	Methods				
	PSO-SVM	Grid-SVM	GA-SVM	BFO-SVM	FOA-SVM
ACC	0.7650 ± 0.0039	0.7648 ± 0.0029	0.7626 ± 0.0053	0.7647 ± 0.0049	0.7746 ± 0.0026
AUC	0.7146 ± 0.0046	0.7119 ± 0.0032	0.7114 ± 0.0062	0.7121 ± 0.0074	0.7234 ± 0.0045
Sensitivity	0.5418 ± 0.0135	0.5359 ± 0.0082	0.5412 ± 0.0113	0.5382 ± 0.0131	0.5507 ± 0.0121
Specificity	0.8874 ± 0.0071	0.8880 ± 0.0058	0.8816 ± 0.0074	0.8861 ± 0.0067	0.8962 ± 0.0040

```
Accuracy: 0.7696 ± 0.0449
AUC: 0.7628 ± 0.2071
Sens: 0.5563 ± 0.0980
Spec: 0.8840 ± 0.0468
```

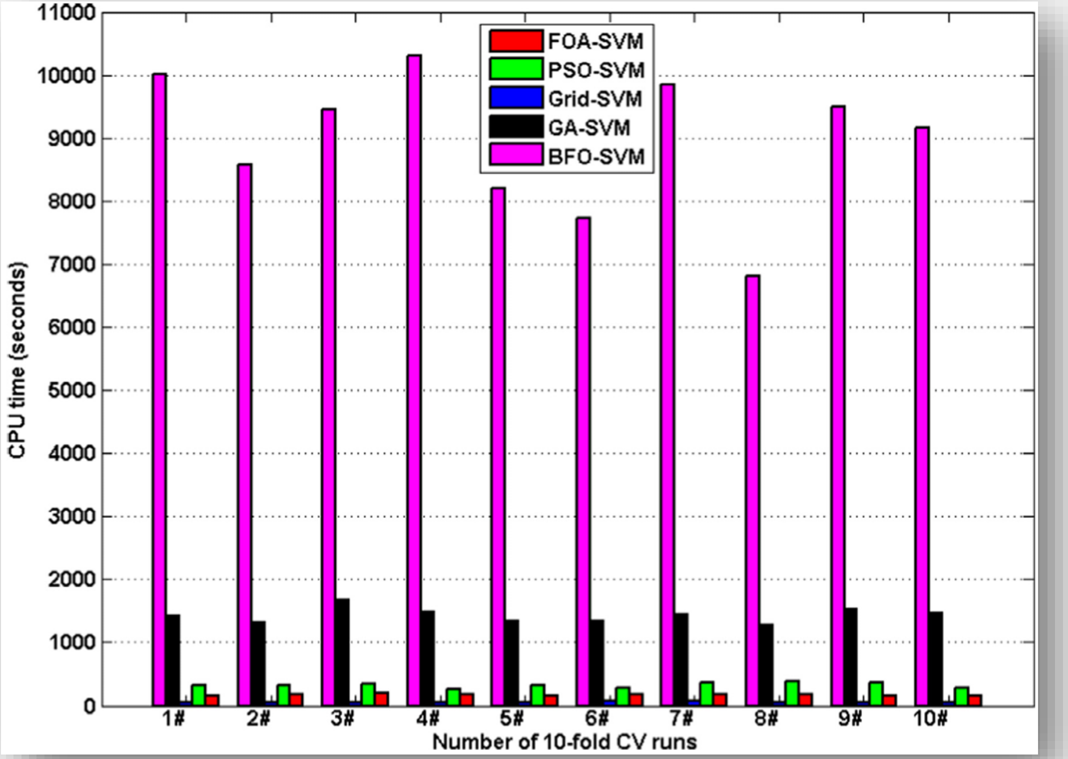
- **Accuracy Across 10 Runs**

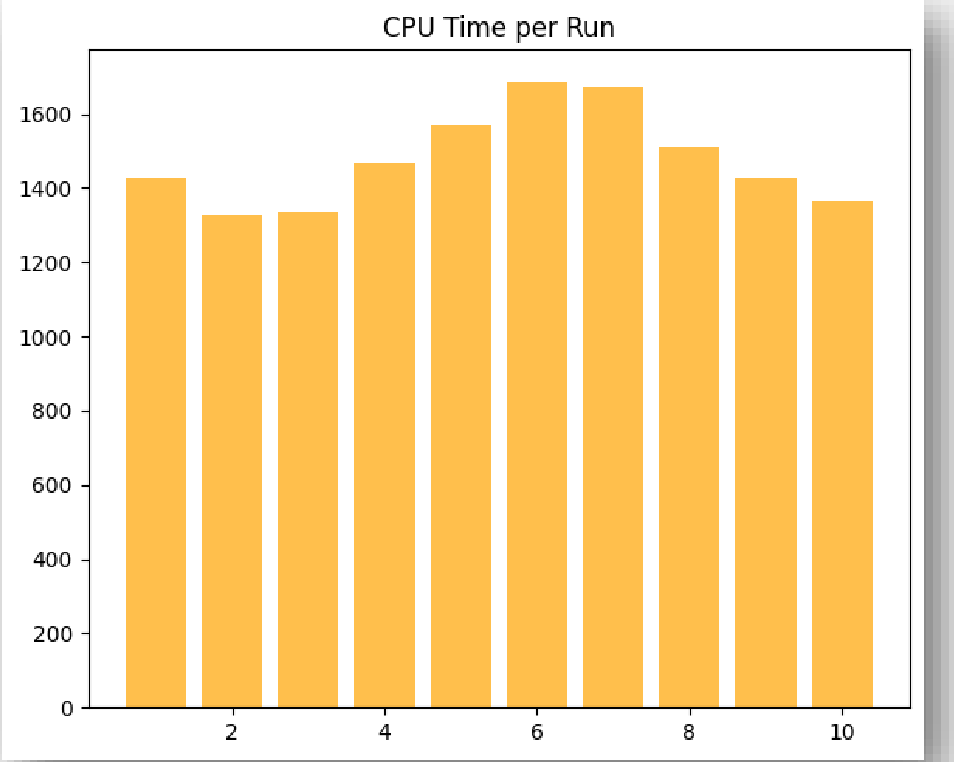




دانشگاه صنعتی امیرکبیر

- CPU Time Comparison – Pima Indians diabetes Dataset





3. Parkinson’s disease diagnosis problem

- 195 instances (23 PD patients + 8 healthy controls) and 22 attributes
- Each subject produced ~6 vowel phonations, each lasting 36 seconds

Label	Attribute	Description
F1	MDVP:Fo (Hz)	Average vocal fundamental frequency
F2	MDVP:Fhi (Hz)	Maximum vocal fundamental frequency
F3	MDVP:Flo (Hz)	Minimum vocal fundamental frequency
F4	MDVP:Jitter (%)	Several measures of variation in fundamental frequency
F5	MDVP:Jitter (Abs)	
F6	MDVP:RAP	
F7	MDVP:PPQ	
F8	Jitter:DDP	
F9	MDVP:Shimmer	Several measures of variation in amplitude
F10	MDVP:Shimmer (dB)	
F11	Shimmer:APQ3	
F12	Shimmer:APQ5	
F13	MDVP:APQ	
F14	Shimmer:DDA	
F15	NHR	Two measures of ratio of noise to tonal components in the voice
F16	HNR	
F17	RPDE	Two nonlinear dynamical complexity measures
F18	D2	
F19	DFA	Signal fractal scaling exponent
F20	Spread1	Three nonlinear measures of fundamental frequency variation
F21	Spread2	
F22	PPE	

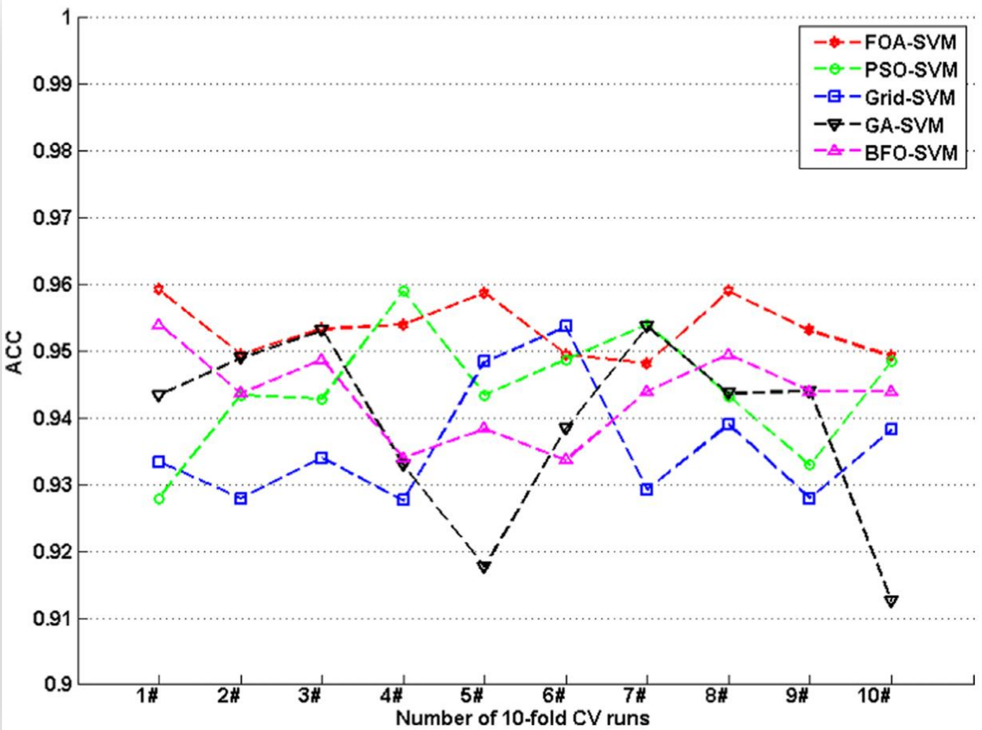
Metrics	*t*-value (significance)			
	PSO-SVM	Grid-SVM	GA-SVM	BFO-SVM
ACC	**2.397(0.040)**	**5.846(0.000)**	3.023(0.014)	**4.953(0.001)**
AUC	0.703(0.500)	**3.902(0.004)**	1.204(0.259)	0.886(0.399)
Sensitivity	**6.047(0.000)**	**8.009(0.000)**	**7.023(0.000)**	**6.873(0.000)**
Specificity	−.741(0.478)	2.192(0.056)	−.317(0.759)	−1.069(0.313)

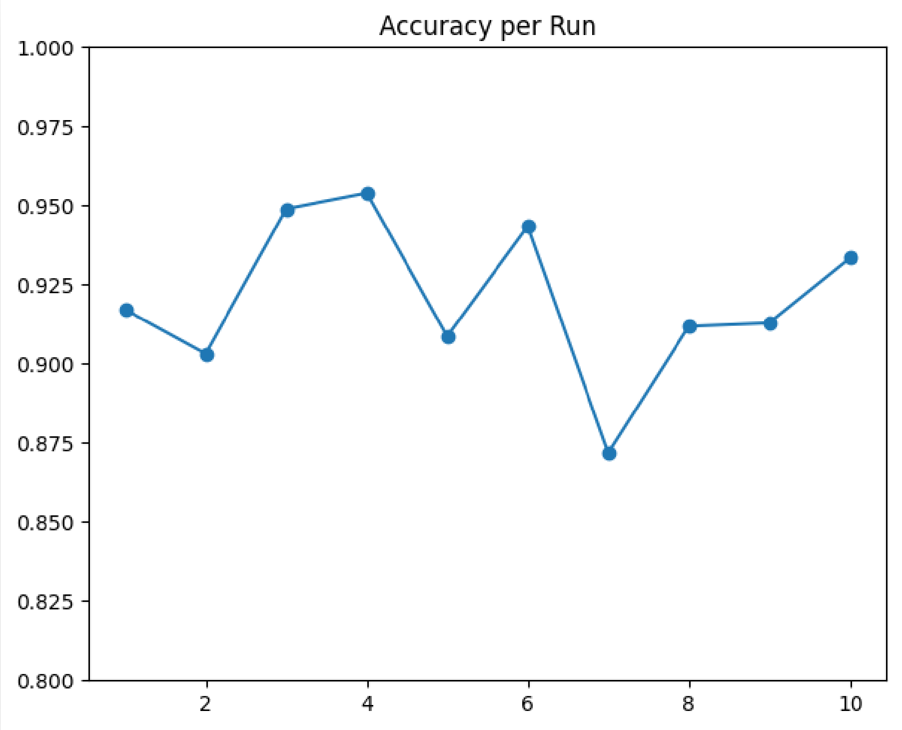
- ## Performance Comparison on Parkinson Dataset

Metrics	Methods				
	PSO-SVM	Grid-SVM	GA-SVM	BFO-SVM	FOA-SVM
ACC	0.9627 ± 0.0038	0.9624 ± 0.0025	0.9564 ± 0.0054	0.9557 ± 0.0056	0.9690 ± 0.0010
AUC	0.9626 ± 0.0042	0.9602 ± 0.0032	0.9529 ± 0.0063	0.9609 ± 0.0044	0.9687 ± 0.0009
Sensitivity	0.9624 ± 0.0070	0.9662 ± 0.0024	0.9627 ± 0.0050	0.9457 ± 0.0082	0.9686 ± 0.0014
Specificity	0.9659 ± 0.0068	0.9545 ± 0.0060	0.9432 ± 0.0099	0.9761 ± 0.0059	0.9689 ± 0.0018

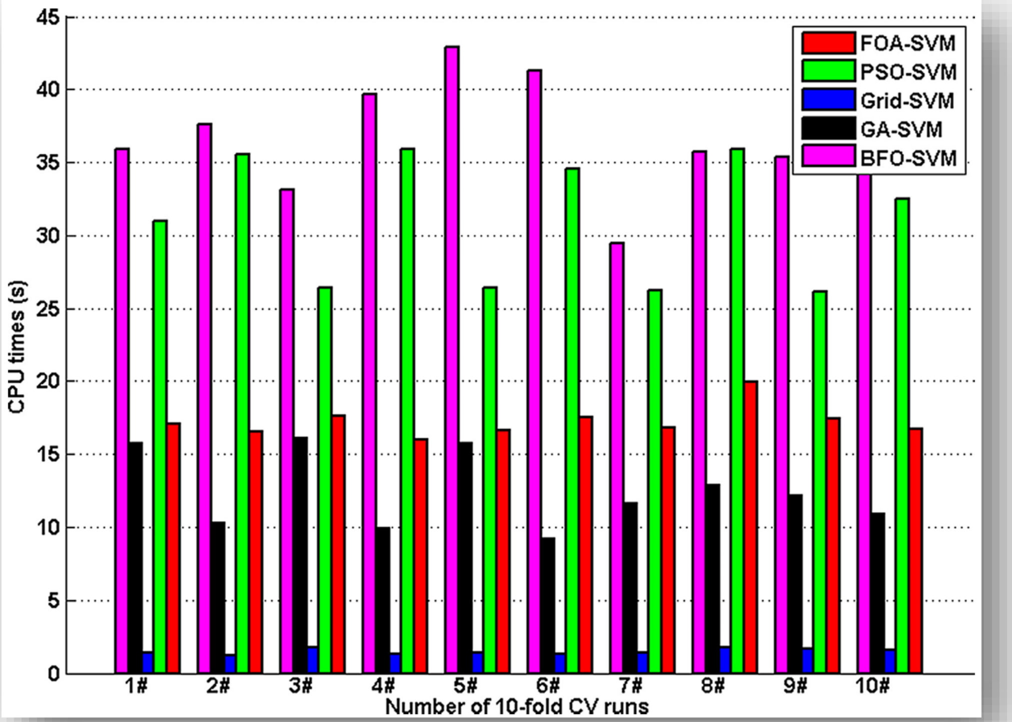
```
Accuracy: 0.9205 ± 0.0627
AUC: 0.9530 ± 0.0697
Sens: 0.9662 ± 0.0499
Spec: 0.7810 ± 0.2164
```

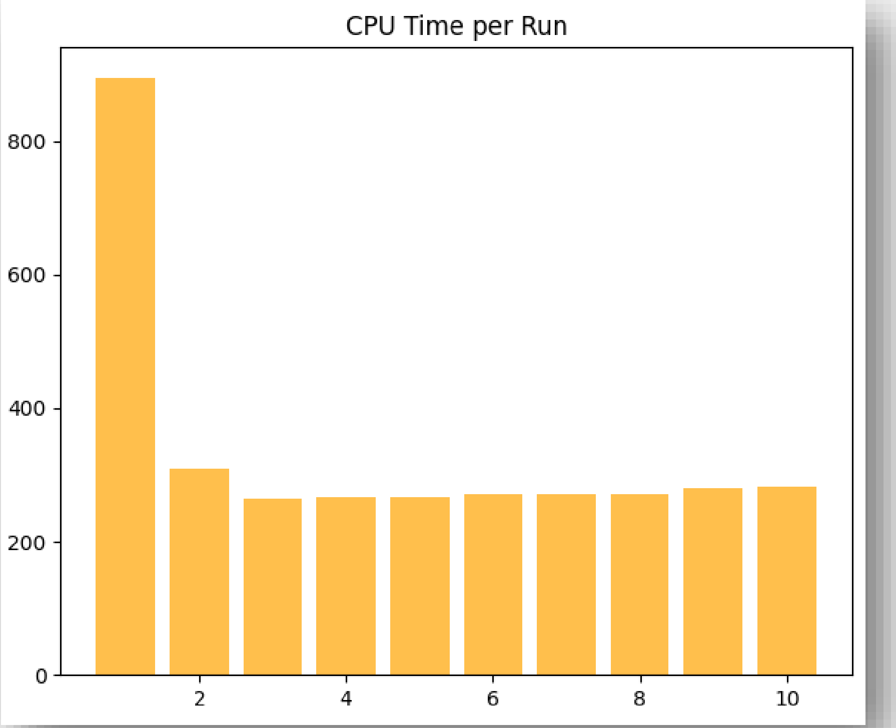
- Accuracy Across 10 Runs





- CPU Time Comparison – Parkinson Dataset





➢ 4. Thyroid disease diagnosis problem

- 215 instances and 5 attributes:
- Class 1: euthyroidism (150) , Class 2: hyper (35) , Class 3: hypo (30)

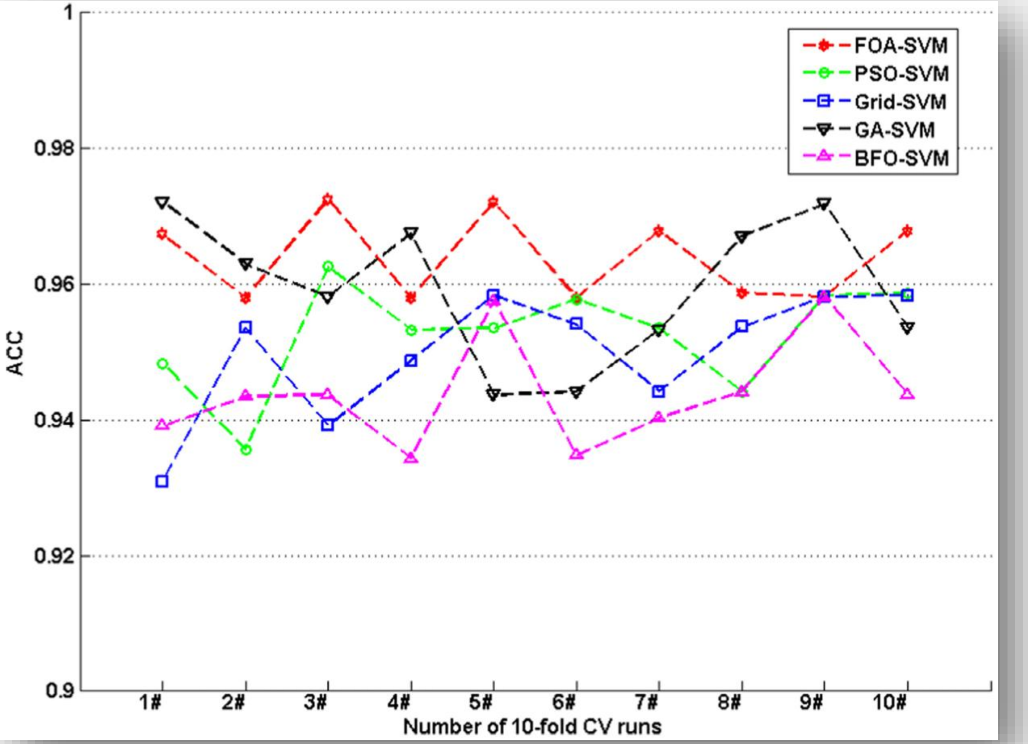
Attributes	Description
F_1	T3-resin uptake test (a percentage).
F_2	Total serum thyroxin as measured by the isotopic displacement method.
F_3	Total serum triiodothyronine as measured by radioimmunoassay.
F_4	Basal thyroid-stimulating hormone (TSH) as measured by radioimmunoassay.
F_5	Maximal absolute difference of TSH value after injection of 200 mg of thyrotropin-releasing hormone as compared to the basal value.

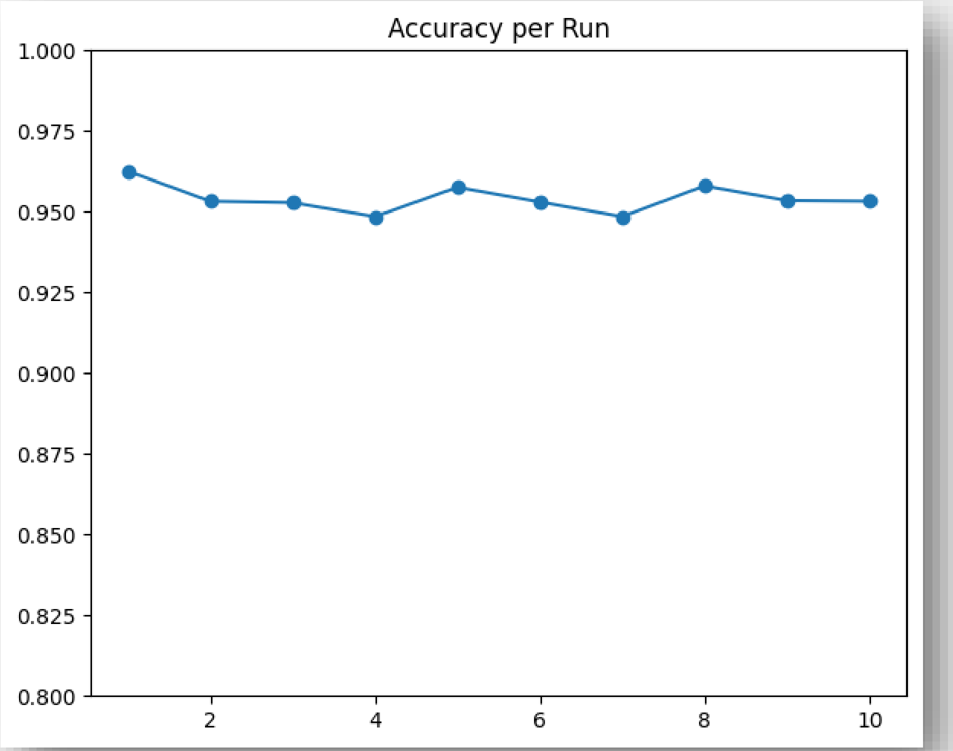
Metrics	Methods				
	PSO-SVM	Grid-SVM	GA-SVM	BFO-SVM	FOA-SVM
ACC	0.9526 ± 0.0080	0.9499 ± 0.0092	0.9594 ± 0.0106	0.9440 ± 0.0082	0.9638 ± 0.0062

Accuracy: 0.9541 ± 0.0409

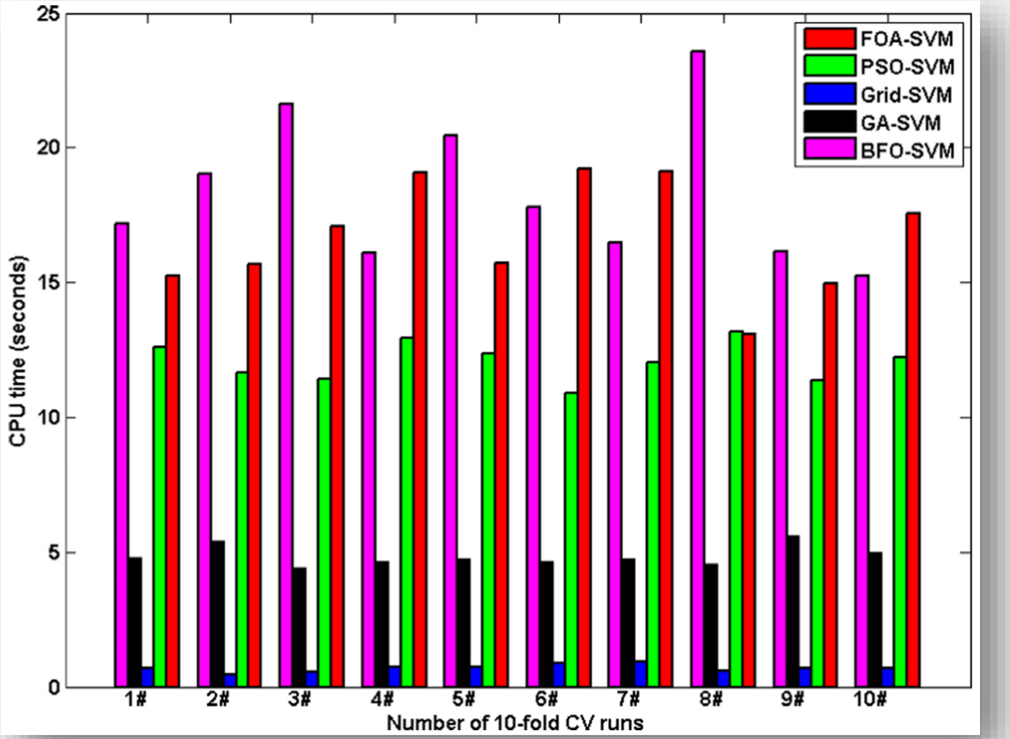
دانشگاه صنعتی امیرکبیر

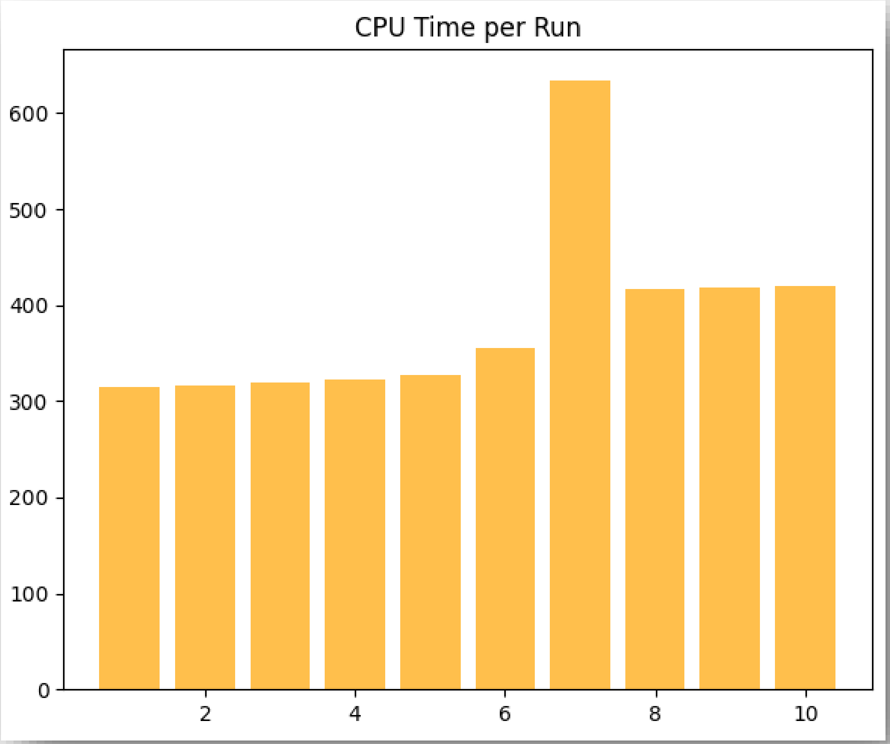
- ## Accuracy Across 10 Runs





- CPU Time Comparison – Thyroid Dataset





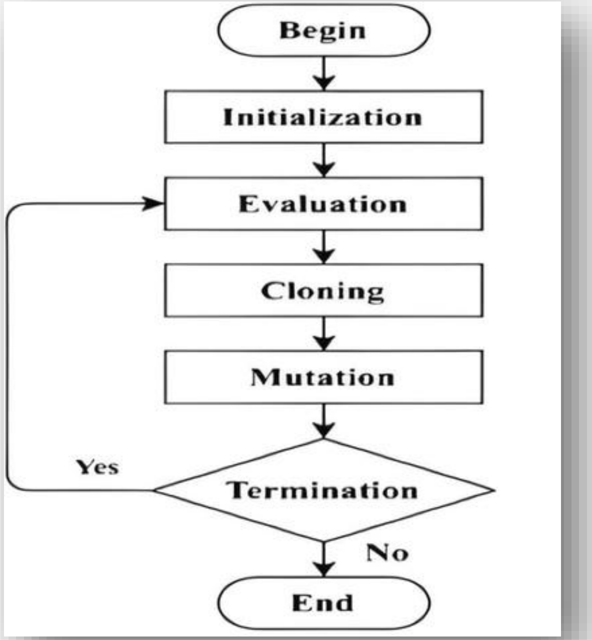
- FOA-SVM achieves higher accuracy and more stable performance than other methods.
- Significantly lower CPU time compared to PSO, GA, and BFO.
- FOA provides an effective and simple approach for SVM parameter tuning.
- A strong candidate for medical data classification tasks.

Improvement

➢ **Clonal Selection Algorithm (CSA)**

- CSA is an immune-inspired population-based optimization algorithm
- Candidate solutions (antibodies) are evaluated using an affinity measure
- High-affinity individuals are selected and cloned
- Adaptive mutation is applied with inverse relation to affinity

$$\text{Mutation} \propto 1/\text{Affinity}$$



Improvement

- <u>Breast cancer diagnosis problem</u>

	FOA-SVM	Implemented FOA-SVM	Implemented CSA-SVM
ACC	0.9690 ± 0.0010	0.9631 ± 0.0224	0.9637 ± 0.0218
AUC	0.9687 ± 0.0009	0.9939 ± 0.0078	0.9938 ± 0.0081
Sensitivity	0.9686 ± 0.0014	0.9461 ± 0.0461	0.9527 ± 0.0385
Specificity	0.9689 ± 0.0018	0.9721 ± 0.0284	0.9695 ± 0.0285

Improvement

- Diabetes disease diagnosis problem

	FOA-SVM	Implemented FOA-SVM	Implemented CSA-SVM
ACC	0.7746 ± 0.0026	0.7696 ± 0.0217	0.7663 ± 0.0459
AUC	0.7234 ± 0.0045	0.7628 ± 0.0207	0.7630 ± 0.0205
Sensitivity	0.5507 ± 0.0121	0.5563 ± 0.0980	0.5469 ± 0.0993
Specificity	0.8962 ± 0.0040	0.8840 ± 0.0468	0.8838 ± 0.0479

Improvement

- **Parkinson's disease diagnosis problem**

	FOA-SVM	Implemented FOA-SVM	Implemented CSA-SVM
ACC	0.9690 ± 0.0010	0.9205 ± 0.0627	0.9477 ± 0.0404
AUC	0.9687 ± 0.0009	0.9530 ± 0.0697	0.9844 ± 0.0251
Sensitivity	0.9686 ± 0.0014	0.9662 ± 0.0499	0.9722 ± 0.0420
Specificity	0.9689 ± 0.0018	0.9716 ± 0.2164	0.8720 ± 0.1312

- **Thyroid disease diagnosis problem**

	FOA-SVM	Implemented FOA-SVM	Implemented CSA-SVM
ACC	0.9638 ± 0.0062	0.9541 ± 0.0409	0.9588 ± 0.0382

References

1) Shen, Liming & Chen, Huiling & Yu, Zhe & Kang, Wenchang & Zhang, Bingyu & Li, Huaizhong & Yang, Bo & Liu, Dayou. "Evolving support vector machines using fruit fly optimization for medical data classification". Knowledge-Based Systems, 2016.

2) V.N. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.

3) W.-T. Pan, A new fruit fly optimization algorithm: Taking the financial distress model as an example, Knowl. Based Syst. 26 (2012) 69–74.

THANK YOU

Any Questions?