Chapter 7

Conclusion and Future Directions

In the previous chapters successful attempts to impart robustness and sparsity to the SVM were presented. In this chapter, the significant contributions of this thesis are summarized. The possible directions for future work are also discussed.

7.1 Conclusion

This thesis mainly focused on robust supervised and semi-supervised machine learning using SVM as the base model. Under supervised machine learning, robust versions of SVM, TWSVM and TSVR were presented. Under semi-supervised machine learning, the robust variant of TSVM was under focus.

In the first contribution, reported in Chapter 3, SVM was made robust against label noise. In that work, a rescaled α-hinge loss function was used in place of the conventional hinge loss in SVM. A non-smooth regularizer was also added with the rescaled α-hinge loss function. The resultant non-smooth objective function was optimized using PDProx dual algorithm. The rescaled α-hinge loss function added robustness against label noise. While, the use of non-smooth regularizer, $\|w\|_2$, added sparsity to the model which leads to improvement in the generalization capability of the model. The proposed formulation was tested on various synthetic and real-world data sets. The experimental results proved that the proposed formulation is more robust and sparse as compared to the models reported in the existing literature on robust SVM. The rate of convergence and the complexity analysis were also provided. The rate of convergence of RSVM-PDProx is $O(1/T)$ where T denotes the number of iterations and the time complexity is $O(n^3)$ where n represents the number of instances in a data set.

In the next contribution, reported in Chapter 4, the regression variant of SVM, TSVR was made robust against the Gaussian and uniform noise in the data sets. In that work, rescaled hinge loss function was used for the first time in a regression framework. This function was used in the conventional TSVR. To optimize the resultant objective function, the half-quadratic optimization technique was used. Experiments were performed in that work to prove the superiority of the proposed approach over the existing methods. The convergence analysis was also given in this chapter. An interesting relationship between the rescaling parameter and the computational time was also observed. As the rescaling parameter increases, the computational time decreases. Although the proposed model enhanced the robustness of TSVR, a limitation was observed with the proposed approach. Since the alternating minimization technique was used in this work, the iterative algorithm approaches a high accuracy solution slowly. However, it is excellent for quickly finding a robust approximate solution.

In another contribution, reported in Chapter 5, TWSVM was used for an important real-world application. In that work, robust TWSVM was used for the diabetic retinopathy detection using eye fundus images. For the DR detection, SVM, the conventional TWSVM and the TWSVM with pinball loss function were used. From the experimental results, it can be concluded that TSVM with pinball loss function performed better DR detection. The computation time corresponding to TWSVM with pinball loss function was also less as compared to the other approaches.

Although the pinball loss function imparted robustness but the model lost its sparseness. To overcome this limitation, in Chapter 6, truncated pinball loss function was proposed [57] in a classification framework. The first two contributions were based on supervised machine learning. Next, a robust semi-supervised machine learning model was also introduced in Chapter 6. This time, a truncated pinball loss function was used with TSVM to impart both the robustness and sparseness in the model. The resultant objective function was optimized in two ways: SGD and the dual problem solver, *mlcv_quadprog*. CCCP was used first on the primal objective function and then SGD was used as the optimization technique. Experiments were performed on both small and large real-world data sets to prove its applicability to a wide area of applications. The time complexity for both the optimization techniques is $O(n^3)$.

The robust formulation was successfully applied to the detection of COVID-19 infected patients using the chest X-ray images. To extract the features, VGG19 was used in this work. The proposed model also helped in finding the labels of the unlabeled samples in the data set.

Next, the possible future scope of the above-discussed work is discussed .

7.2 Future Directions

The above-discussed contributions in the thesis may lead to several future research directions:

(i) In this thesis, SVM and its variants were made robust against label noise in the data set. To achieve a robust model, robust loss functions were used. This can be extended to other machine learning models as well. It will be interesting to add robustness to other machine learning models by using various robust loss functions in their formulation and solving the corresponding optimization problems.

(ii) In the above-discussed contributions, supervised and semi-supervised variants of SVM were considered for adding robustness against noise in the data set. In the future, it can be extended to the unsupervised variant of SVM as well.

(iii) Deep learning is an active research field of machine learning that gives promising results in many of the application areas like image processing, video processing, speech recognition etc. The proposed loss functions can be applied to the deep learning models as well.

(iv) The proposed formulations can be applied to other real-world applications like in the field of medical diagnosis (classification) and statistical arbitrage (regression), etc.