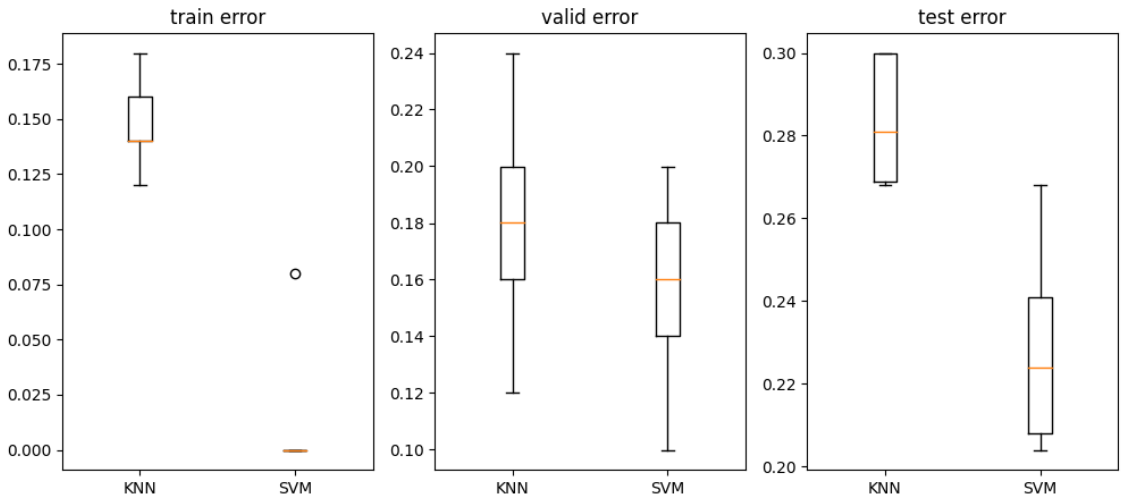
Predictive learning HW1

P1

(a)

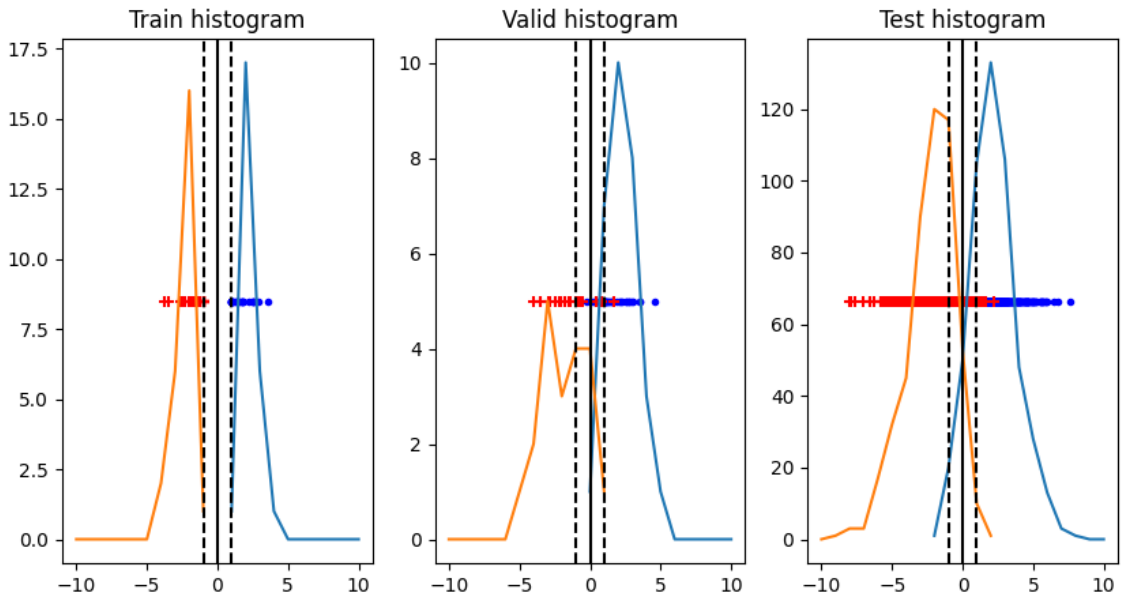
I follow the question, and training with 5 time for each model with validation dataset. KNN model tune k neihbor, and SVM tune c. Recording the training, validation, testing error with boxplot.



We can find SVM has better performence than KNN algorithm on this dataset. So SVM would be ideal model than KNN on this dataset.

(b)

According to question, I draw histogram-of-projections with SVM model.



By observation, we can find the SV is not large. So the first bound will be better performence on this dataset.

P2

(a)

Following the question, I get the MSE on validation dataset.

SVM MSE on validation set

	ε=0	ε=2	ε=4	ε=6	ε=8
γ=2^-5	194.36116904729616	196.46571793352155	187.90881360456532	188.03397368688252	175.86522300206335
γ=2^-4	146.0941092407987	137.8862122577435	128.72261038712756	118.84028471604924	121.82865313773286
γ=2^-3	110.59055605907068	111.38178928869097	113.39622721718528	107.50774378624658	108.98400965465238
γ=2^-2	127.61112184344009	120.8385266150174	115.17186726902128	124.58910970114206	133.48078633518472
γ=2^-1	167.79822304640683	171.65269796799706	182.72855611399845	192.34365861171648	203.27123071246808
γ=2^0	279.4949372642348	290.75010764196645	302.0882262275319	313.33690066548564	325.51366538161284
γ=2^1	436.573167394253	441.71306552067216	447.5912825678151	454.04467484082403	461.02961507038896
γ=2^2	514.2476763210298	514.0704887115207	514.3307719084924	515.2781829378936	516.9047879403072
γ=2^3	531.0863375251217	529.5669508957046	528.4845384985234	528.1363290744217	528.6192547300057
γ=2^4	534.5380201637058	532.7271826820659	531.3601565238725	530.7385518059232	531.0026067301593
γ=2^5	534.8539891896997	533.0179401700171	531.6267125208363	530.981819946823	531.2277310624156

According to validation MSE. We selection SVM model with epsilon = 6 and gamma = 2^-3. With this model, the MSE with test dataset is 101.2220405004163 and NRMSEE is 0.3261091899071216.

(b)

I use PPR model with same training dataset, and tune the model compleity with r (The number of terms in the underlying additive model). And record it MSE on vaildation set.

PPR MSE on validation set

	mse
r=1	79.107891790066
r=2	99.2883682602048
r=3	78.45575789561812
r=4	94.40333786034233
r=5	125.80046800145199
r=6	135.39768513252878
r=7	143.6596061742737
r=8	159.77411219676588
r=9	187.9439442532891
r=10	198.8971555564935
r=11	282.24218848294544
r=12	238.9831112064941
r=13	368.0388571810513
r=14	396.9673684947475
r=15	348.1965694907426
r=16	503.7307414227308
r=17	422.0144171417416
r=18	487.2846088532354
r=19	488.9702644031849
r=20	589.5714279698271

I found the r = 3 is the best model, and the MSE on test dataset with best PPR model is 71.14189790761668 and NRMSE is 0.2721736435450077.

Test error on two model:

	MSE	**NRMSE**
SVM	101.2220405004163	0.3261091899071216
PPR	**71.14189790761668**	**0.2721736435450077**

According to table. The PPR model performence is obviously better than SVM model. SVMs might not be effective for data sets which relationships between variables are more complex.