Select

$$\min_{Q,f} \pi R(P \cup Q_P, +) + R(U, -) - \pi R(P \cup Q_P, -)$$

$$= \min_{Q,f} \pi \left[R(P \cup Q_P, +) - R(P \cup Q_P, -) \right] + R(U, -) \qquad S = U \setminus Q$$

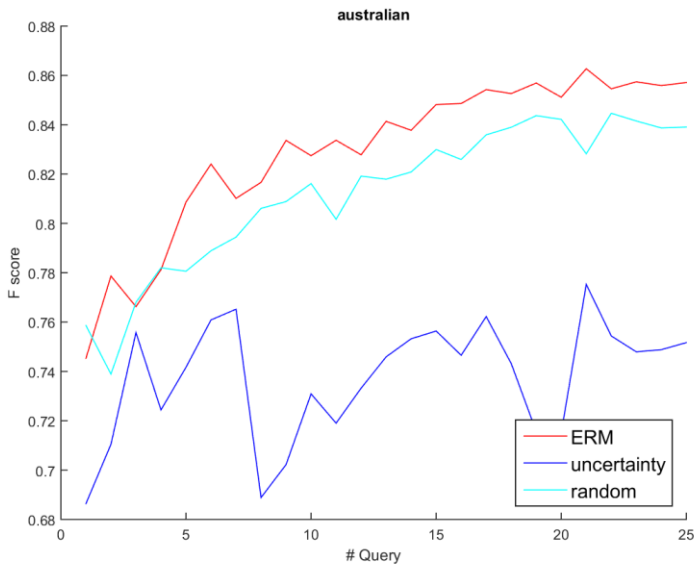
$$= \min_{Q,f} \pi \left[\frac{1}{|P \cup Q_P|} \sum_{x_i \in P \cup Q_P} l(f(x_i), +1) \right] + \frac{1}{|U|} \sum_{x_j \in U} l(f(x_j), -1) - \pi \left[\frac{1}{|P \cup Q_P|} \sum_{x_i \in P \cup Q_P} l(f(x_i), -1) \right]$$

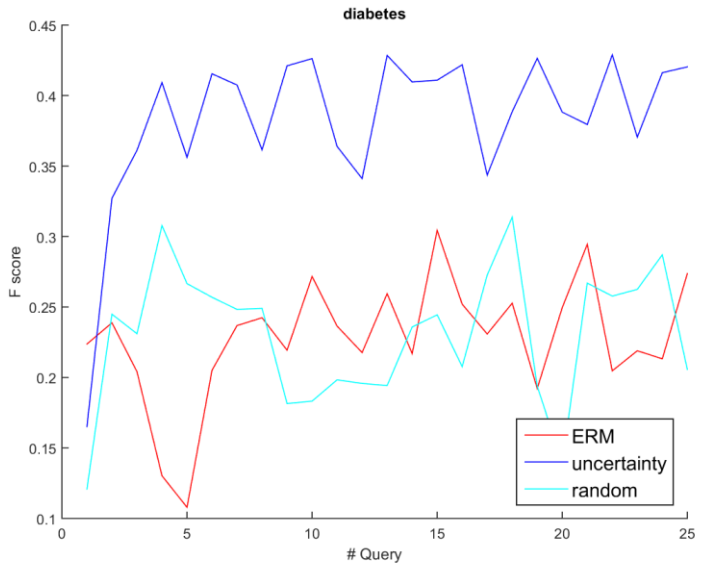
$$= \min_{\beta^T 1_S = b, f} \frac{\pi}{n_P + b} \left[\sum_{x_i \in P} l(f(x_i), +1) - l(f(x_i), -1) + \sum_{x_i \in S} \beta_i \alpha_i \left[l(f(x_i), +1) - l(f(x_i), -1) \right] \right] + \frac{1}{n_U} \sum_{x_j \in U} l(f(x_j), -1)$$

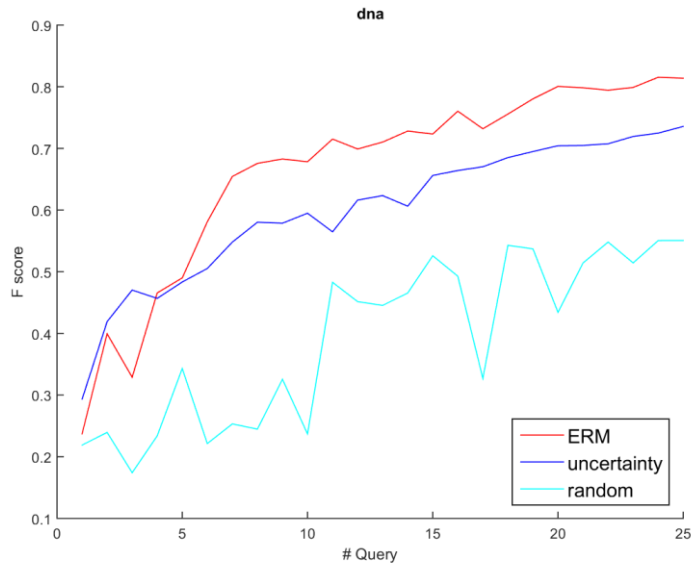
Datasets

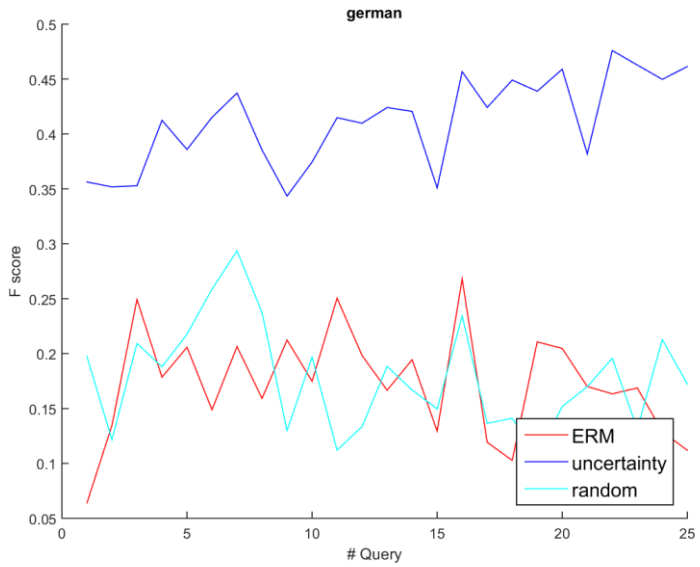
datasets	# train	# test	# features	# labels
australian	483	207	42	2
diabetes	538	230	8	2
dna	1400	1186	180	3
german	700	300	59	2
krvskp	2237	959	36	2
letter	10500	5000	16	26
pendigits	7494	3498	16	10
protein	14895	6621	357	3
satimage	3104	2000	36	6
sick	2641	1131	31	2
svmguide1	3089	4000	4	2
usps	7291	2007	256	10

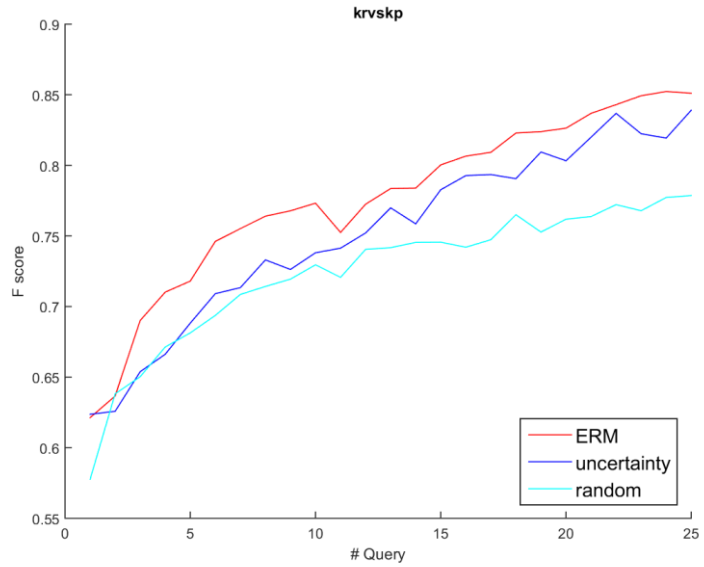
Results

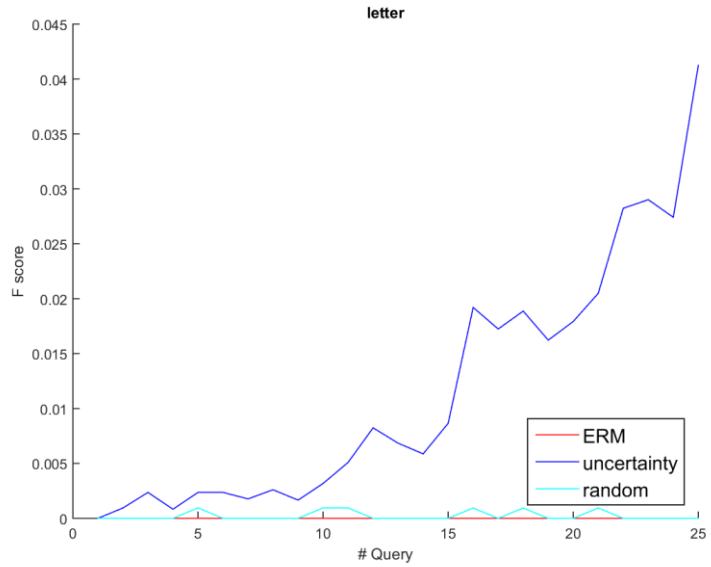

australian
F score
Query
ERM
uncertainty
random

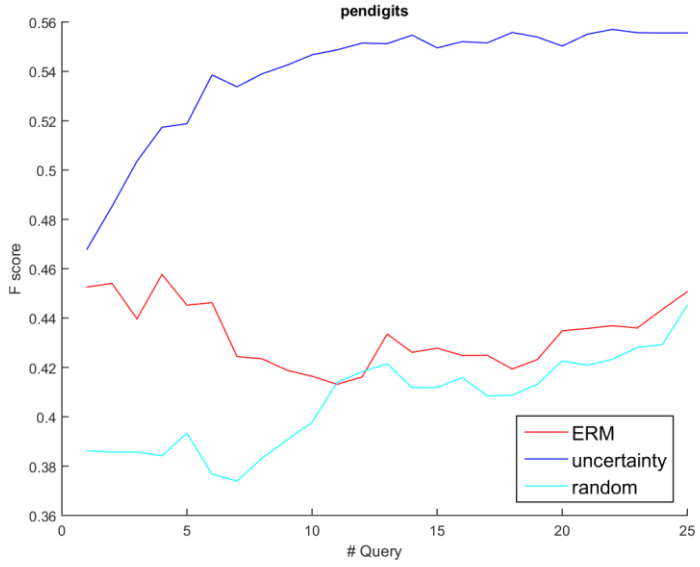

diabetes
F score
Query
ERM
uncertainty
random

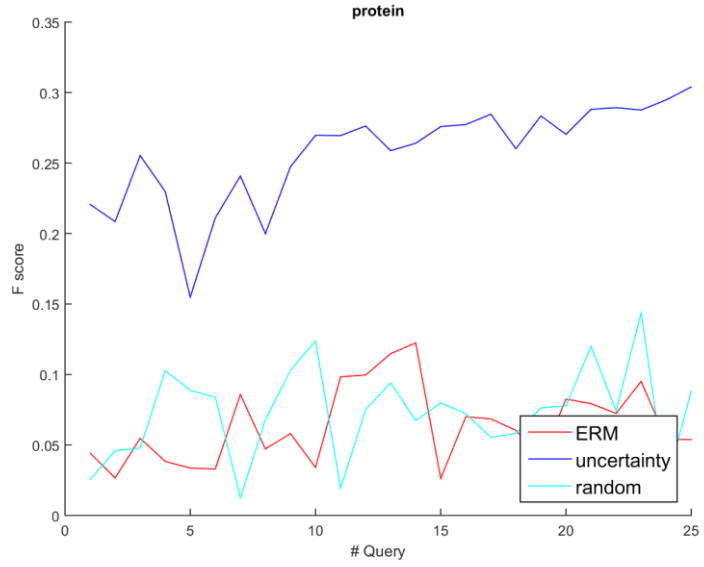

dna
F score
Query
ERM
uncertainty
random

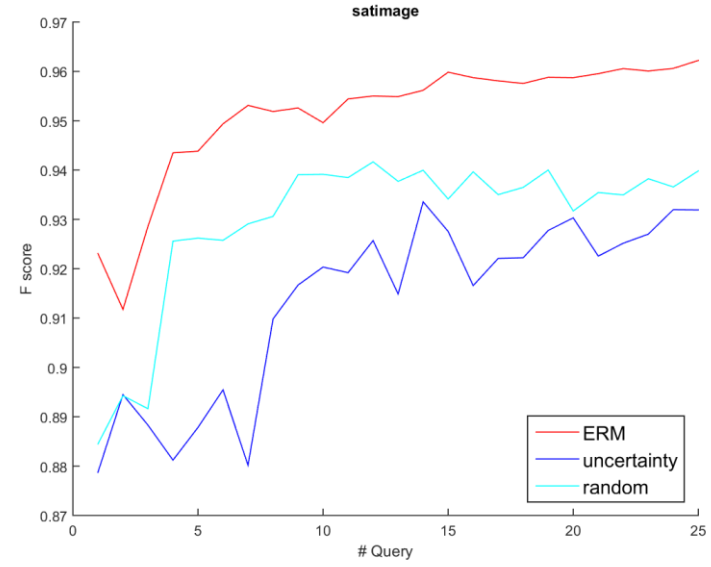

german
F score
Query
ERM
uncertainty
random

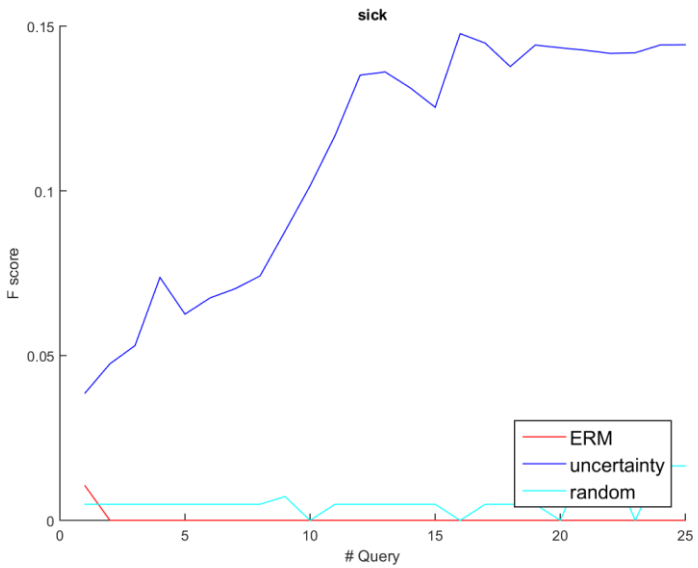

krvskp
F score
Query
ERM
uncertainty
random

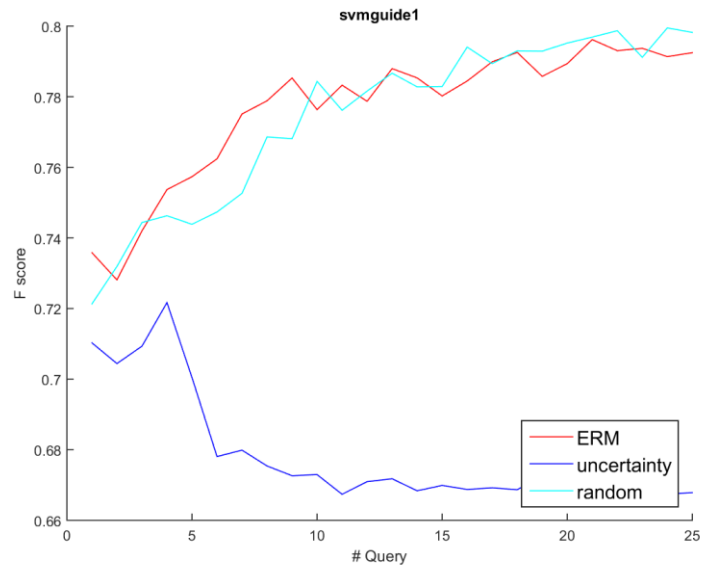

letter
F score
Query
ERM
uncertainty
random

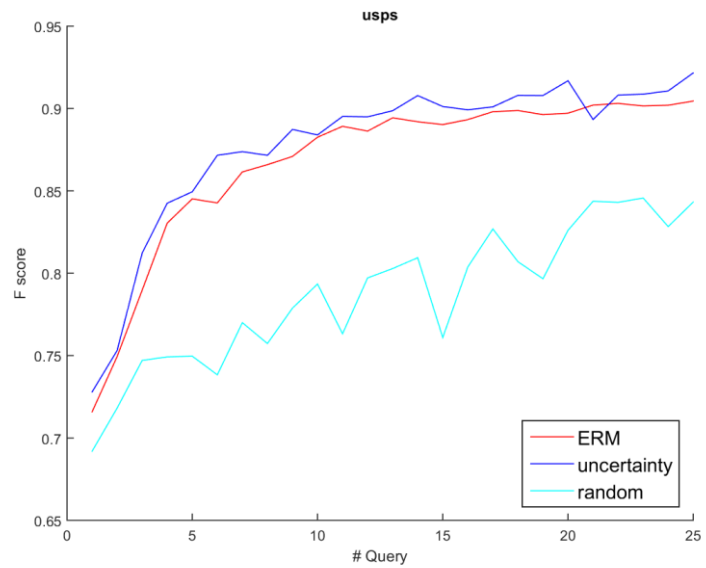

pendigits
F score
Query
ERM
uncertainty
random


protein
F score
Query
ERM
uncertainty
random


satimage
F score
Query
ERM
uncertainty
random


sick
F score
Query
ERM
uncertainty
random


svmguide1
F score
Query
ERM
uncertainty
random


usps
F score
Query
ERM
uncertainty
random

Problem

(1) $$\min \sum_{x_i \in S} \beta_i \alpha_i l(f(x_i), +1) - \beta_i \alpha_i l(f(x_i), -1)$$

$$f(x_i) \approx 1$$

(2) Convergence:

$$|\text{prev\_risk} - \text{post\_risk}| < \varepsilon$$