	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Posts	Visits
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 5$	8.441	0.922	0.064	0.493	1601.321	33.000	545.371
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 10$	8.632	0.924	0.064	0.494	1593.763	33.000	545.206
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 15$	8.913	0.924	0.064	0.494	1594.276	33.000	545.381
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 20$	9.382	0.923	0.063	0.493	1597.533	33.000	545.495
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 25$	9.905	0.927	0.063	0.495	1597.295	33.000	545.619
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 30$	10.836	0.925	0.063	0.494	1587.734	33.000	545.619
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 35$	12.044	0.927	0.064	0.495	1608.698	33.000	545.722
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 40$	12.400	0.927	0.063	0.495	1593.062	33.000	545.649
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 45$	12.462	0.928	0.063	0.496	1587.913	33.000	545.649
$w = 5, \mathbf{v}, \lvert\mathbf{v}\rvert = 50$	12.994	0.926	0.063	0.495	1581.241	33.000	545.804

Table 1: Experiment results: Varying vocabulary size

1 Method

We extracted the timestamp, author and text content for each post in each thread from the forums. Our dataset was obtained from from `avsforums.com`.

Previous time differences All the time differences between posts made in the window. ($\mathbf{t}_\Delta$)

Time-based features Day of week, Hour of day. Provides contextual information about when the post was made. ($\mathbf{t}_{\text{ctx}}$)

Content features (text) Word frequency counts are used for this set of experiments. Using regression, we find the top K variables that the actual Δ_t depends on. Table 1 reflect the results of the experiments done with varying values of K.

1.1 Potential errors

To be thorough, let us also enumerate the types of errors that a model making predictions could encounter.

The model can potentially make a prediction such that the next visit comes before the arrival of the next post. The predictions being made are the Δ_t between the posts, rather than the visitation times, hence, it is possible for the model to make a prediciton that occurs before the current time. An erroneous prediction can also cause the crawler to come in before the next post (two, or more, visits, but nothing new fetched). Errors of this type waste bandwidth, since the crawler will make an unneccessary visit to the page.

Another type of error would have the prediction causing the next visit to come some time after a post. Since most predictions are almost never fully accurate, there will be some time between the post is made and hte page is fetched. These errors are still relatively acceptable, but the time difference between the post arriving and the visit should be minimised. The visit could also come more than one post later. Errors of this kind incur a penalty on the freshness of the data, more so than the after one post, especially if the multiple posts are far apart time-wise.

In the following experiments, the threads chosen from our extracted dataset are those with a 100 to 1000 posts. This amounted to 97 threads. The first 75% of the thread was used as training data, while the remaining 25% was used as test data. We used Support Vector machines for this regression task, employing a Radial Basis Function kernel as our learning algorithm.

The SVR module from the Python library scikit-learn was used in the implementation of this experiment.

1.2 Evaluation metrics

We use *Mean Absolute Percentage Error* (MAPE), to measure the performance of the learnt model. This value is given by

$$\frac{1}{N}\sum_{i=1}^{N}\left|\frac{A_i - F_i}{A_i}\right|$$

where A_i is the actual value, and F_i is the forecasted value for the instance i. Realistically, the model would not be able to come into contact with every possible window, since chances are it will make an error that causes it to visit a thread late, causing it to miss two posts or more. This value does not reflect how well the model will do in a real-time setting, but gives an idea of how far off the model is given a window.

We also want to know the *timeliness* of the model's visits. Yang et. al. [**?**] has a metric for measuring this. Taking Δt_i as the time difference between a post i and it's download time, the timeliness of the algorithm is given by

$$T = \frac{1}{N}\sum_{i=1}^{N} \Delta t_i$$

A good algorithm would give a low T-score. However, a crawler that hits the site repeatedly performs well according to this metric. The authors account for this by setting a bandwidth (fixed number of pages per day) for each iteration of their testing. In our experimental results, we also take into account the number of page requests made in comparison to the number of posts.

$$\begin{aligned} Pr_{miss} &= \frac{\sum_{i=1}^{N-k}\left[\Theta_{ref\_hyp}(i,k)\right]}{\sum_{i=1}^{N-k}\left[\Delta_{ref}(i,k)\right]} \\ Pr_{fa} &= \frac{\sum_{i=1}^{N-k}\left[\Psi_{ref\_hyp}(i,k)\right]}{N-k} \end{aligned} \qquad \begin{aligned} \Delta_{ref}(i,k) &= \begin{cases} 1, & \text{if } r(i,k) > 0 \\ 0, & \text{otherwise} \end{cases} \\ \Theta_{ref\_hyp}(i,k) &= \begin{cases} 1, & \text{if ends with post} \\ 0, & \text{otherwise} \end{cases} \\ \Psi_{ref\_hyp}(i,k) &= \begin{cases} 1, & \text{if ends with visit} \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

Split the Pr_{error} measure into 3 parts:

Probability of having visits before the first post in the window. Probability of having more visits than posts after the first post. Probability of having more posts than visits after the first post.

Viewing the posts made during the thread's lifetime as segmentations of the thread, and the visits made as hypotheses of where the segmentations are, we use the Pr_{error} metric from Georgescul et. al., 2006 as a measure of how close the predictions are to the actual posts. An example can be seen in Figure 1.

Function that gives me:

$$\uparrow \quad p_1, \underbrace{p_2, \quad p_3, \quad p_4, \uparrow}_{\text{sliding window}} \uparrow$$

Figure 1: An example of the sliding window metric. The metric is made up of two components: First, the probability that, given at least one post is present in the window, there are more visits than post. Secondly, the probability that there are more posts than visits for a given window. The weighted sum of this gives the overall Pr_{error}

Increase in visit to post ratio increase Increase in interval increase

Increase between post and visit increase (0,∞ or 1)

Increase between visit and visit decrease (∞ or 1,0) Increase between post and post increase (0,∞ or 1)

$$T = \frac{\sum_{e=1}^{|E|-1} \Psi(e_t, e_{t+1})}{|E|-1} \quad \Psi(e_t, e_{t+1}) = \begin{cases} 1 - e^{-(e_{t+1}-e_t)} & \text{if post,visit} \\ e^{-(e_{t+1}-e_t)} & \text{if visit,visit} \end{cases}$$

1.3 Normalising the T-score and Visit/Post ratio

We normalise the T-score to get a comparable metric across all the threads. In order to do this, we consider again the thread posts and visits as a sequence of events. We then define the *lifetime*, denoted as l, of the thread as the time between the first post and the last post. Any visits that occur after the last post are ignored.

We then consider the worst case in terms of timeliness, or misses. This would be the case where the visit comes at the end, at the same time as the post. So we get a value T_{max} and P_{miss} such that

$$P_{\text{miss}} = \frac{T}{T_{\text{max}}} = \frac{N \cdot T}{\sum_p l - p_t}$$

It is difficult to consider the worst case in terms of false alarms, or visits that retrieve nothing. There could be an infinite number of visits made if we are to take the extreme case. In order to get around this, we consider discrete time frames in which a visit can

	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Posts	Visits
Average $w = 5$	330.285	0.951	0.054	0.502	6418.208	33.000	498.742
Average $w = 10$	305.557	0.955	0.053	0.504	4598.955	31.680	497.351
Average Δ_t	174.004	0.938	0.065	0.501	1764.474	34.000	574.031
$w = 5, \mathbf{t}_\Delta$	18.884	0.931	0.064	0.498	1541.595	33.000	547.062
$w = 5, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	18.885	0.931	0.064	0.498	1541.592	33.000	547.062
$w = 5, \mathbf{v}$	9.382	0.923	0.063	0.493	1597.533	33.000	545.495
$w = 5, \mathbf{v}, \mathbf{t}_\Delta$	18.877	0.931	0.064	0.498	1541.588	33.000	547.062

Table 2: Experiment results

occur. Since for this dataset, our time granularity is in terms of minutes, we shall use minutes as our discrete time frame.

With this simplified version of our series of events, we can then imagine a worst-case performing revisit policy that visits at every single time frame. This gives us

$$P_{\text{FA}} = \frac{|V|}{l - |P|}$$

where l is in units of our specified discrete time frame.

2 Results

The results for experiments done with different combinations of the above specified features are shown in Table 2.

Table 3: Experiment results

	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Visit/Post
Average $w = 1$	162.495 ± 30.282	0.958 ± 0.008	0.049 ± 0.010	0.504 ± 0.003	18862.320 ± 4267.812	16.142 ± 7.049
Average $w = 2$	239.105 ± 57.278	0.953 ± 0.009	0.052 ± 0.011	0.503 ± 0.003	11599.851 ± 1906.819	16.199 ± 7.243
Average $w = 3$	282.490 ± 57.885	0.955 ± 0.009	0.052 ± 0.011	0.503 ± 0.003	8334.424 ± 1330.301	16.646 ± 7.388
Average $w = 4$	327.016 ± 59.688	0.954 ± 0.010	0.053 ± 0.011	0.503 ± 0.003	7146.832 ± 1129.593	16.840 ± 7.433
Average $w = 5$	330.285 ± 61.259	0.951 ± 0.010	0.054 ± 0.012	0.502 ± 0.003	6418.208 ± 962.716	16.464 ± 7.386
Average $w = 6$	343.840 ± 61.801	0.953 ± 0.010	0.053 ± 0.011	0.503 ± 0.003	5482.831 ± 797.896	16.924 ± 7.664
Average $w = 7$	331.373 ± 55.988	0.954 ± 0.010	0.053 ± 0.011	0.504 ± 0.003	5056.252 ± 733.048	17.037 ± 7.672
Average $w = 8$	321.779 ± 52.948	0.954 ± 0.010	0.053 ± 0.012	0.504 ± 0.003	4785.875 ± 631.967	17.088 ± 7.651
Average $w = 9$	310.990 ± 50.260	0.954 ± 0.010	0.053 ± 0.012	0.504 ± 0.003	4542.149 ± 560.353	17.058 ± 7.626
Average $w = 10$	305.557 ± 49.021	0.955 ± 0.010	0.053 ± 0.012	0.504 ± 0.003	4598.955 ± 682.458	17.291 ± 7.872
Average $w = 11$	300.200 ± 48.548	0.956 ± 0.010	0.053 ± 0.012	0.505 ± 0.003	4432.093 ± 650.508	17.368 ± 7.827
Average $w = 12$	312.179 ± 55.062	0.955 ± 0.010	0.054 ± 0.012	0.505 ± 0.003	4064.795 ± 593.199	17.890 ± 8.283
Average $w = 13$	302.924 ± 53.841	0.951 ± 0.010	0.055 ± 0.012	0.503 ± 0.003	4023.003 ± 586.345	18.114 ± 8.504
Average $w = 14$	314.973 ± 55.676	0.951 ± 0.011	0.054 ± 0.012	0.503 ± 0.003	3904.468 ± 566.171	18.357 ± 8.782
Average $w = 15$	308.547 ± 53.488	0.954 ± 0.011	0.054 ± 0.012	0.504 ± 0.003	3833.605 ± 600.824	18.337 ± 8.727
Average $w = 16$	298.138 ± 51.469	0.952 ± 0.011	0.054 ± 0.012	0.503 ± 0.003	3558.455 ± 495.589	18.387 ± 8.757
Average $w = 17$	288.236 ± 49.821	0.950 ± 0.011	0.055 ± 0.012	0.502 ± 0.003	3459.059 ± 473.284	18.425 ± 8.880
Average $w = 18$	266.409 ± 42.039	0.950 ± 0.012	0.055 ± 0.012	0.502 ± 0.003	3548.269 ± 477.675	18.098 ± 8.586
Average $w = 19$	264.901 ± 42.287	0.953 ± 0.010	0.054 ± 0.012	0.503 ± 0.003	3650.814 ± 542.684	18.232 ± 8.635
Average $w = 20$	265.124 ± 42.829	0.953 ± 0.010	0.054 ± 0.012	0.504 ± 0.003	3340.929 ± 444.908	18.102 ± 8.541

	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Visit/Post
$w = 1, \mathbf{t}_\Delta$	13.797 ± 1.769	0.929 ± 0.014	0.067 ± 0.015	0.498 ± 0.002	1576.082 ± 253.300	18.267 ± 7.290
$w = 2, \mathbf{t}_\Delta$	18.192 ± 2.096	0.928 ± 0.014	0.066 ± 0.014	0.497 ± 0.002	1550.939 ± 251.635	18.465 ± 7.490
$w = 3, \mathbf{t}_\Delta$	18.747 ± 2.128	0.929 ± 0.014	0.065 ± 0.014	0.497 ± 0.002	1511.769 ± 238.844	18.590 ± 7.582
$w = 4, \mathbf{t}_\Delta$	18.910 ± 2.105	0.929 ± 0.015	0.065 ± 0.014	0.497 ± 0.002	1533.559 ± 241.443	18.524 ± 7.587
$w = 5, \mathbf{t}_\Delta$	18.884 ± 2.121	0.931 ± 0.014	0.064 ± 0.014	0.498 ± 0.002	1541.595 ± 232.272	17.907 ± 7.508
$w = 6, \mathbf{t}_\Delta$	19.269 ± 2.214	0.933 ± 0.014	0.063 ± 0.014	0.498 ± 0.002	1560.129 ± 235.673	18.258 ± 7.769
$w = 7, \mathbf{t}_\Delta$	19.213 ± 2.216	0.933 ± 0.014	0.063 ± 0.014	0.498 ± 0.002	1441.079 ± 200.736	18.294 ± 7.762
$w = 8, \mathbf{t}_\Delta$	19.409 ± 2.254	0.934 ± 0.014	0.062 ± 0.014	0.498 ± 0.002	1414.262 ± 184.353	18.265 ± 7.730
$w = 9, \mathbf{t}_\Delta$	19.504 ± 2.266	0.936 ± 0.013	0.062 ± 0.014	0.499 ± 0.002	1472.411 ± 190.027	18.194 ± 7.701
$w = 10, \mathbf{t}_\Delta$	19.647 ± 2.317	0.937 ± 0.013	0.061 ± 0.014	0.499 ± 0.002	1488.688 ± 196.648	18.371 ± 7.947
$w = 11, \mathbf{t}_\Delta$	19.862 ± 2.358	0.936 ± 0.013	0.061 ± 0.014	0.498 ± 0.002	1403.129 ± 169.845	18.420 ± 7.916
$w = 12, \mathbf{t}_\Delta$	19.934 ± 2.373	0.934 ± 0.014	0.061 ± 0.014	0.497 ± 0.002	1444.162 ± 174.673	18.907 ± 8.372
$w = 13, \mathbf{t}_\Delta$	19.957 ± 2.368	0.936 ± 0.013	0.062 ± 0.014	0.499 ± 0.002	1465.005 ± 187.948	19.086 ± 8.590
$w = 14, \mathbf{t}_\Delta$	20.066 ± 2.423	0.937 ± 0.013	0.061 ± 0.014	0.499 ± 0.002	1514.741 ± 188.504	19.285 ± 8.865
$w = 15, \mathbf{t}_\Delta$	20.195 ± 2.426	0.939 ± 0.013	0.061 ± 0.014	0.500 ± 0.002	1443.138 ± 183.408	19.234 ± 8.805
$w = 16, \mathbf{t}_\Delta$	20.253 ± 2.458	0.939 ± 0.013	0.061 ± 0.014	0.500 ± 0.002	1500.906 ± 199.944	19.259 ± 8.831
$w = 17, \mathbf{t}_\Delta$	20.335 ± 2.492	0.941 ± 0.013	0.061 ± 0.014	0.501 ± 0.001	1578.627 ± 215.405	19.258 ± 8.951
$w = 18, \mathbf{t}_\Delta$	20.049 ± 2.342	0.937 ± 0.014	0.060 ± 0.013	0.499 ± 0.002	1639.302 ± 231.755	18.904 ± 8.652
$w = 19, \mathbf{t}_\Delta$	20.021 ± 2.339	0.936 ± 0.013	0.059 ± 0.013	0.498 ± 0.001	1594.704 ± 213.182	19.030 ± 8.698
$w = 20, \mathbf{t}_\Delta$	20.220 ± 2.374	0.938 ± 0.012	0.059 ± 0.013	0.499 ± 0.001	1584.171 ± 227.209	18.880 ± 8.602
$w = 1, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	15.705 ± 1.880	0.929 ± 0.014	0.067 ± 0.015	0.498 ± 0.001	1562.579 ± 252.719	18.284 ± 7.291
$w = 2, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	18.382 ± 2.108	0.927 ± 0.014	0.066 ± 0.014	0.496 ± 0.002	1549.568 ± 251.698	18.469 ± 7.491
$w = 3, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	18.800 ± 2.130	0.929 ± 0.014	0.065 ± 0.014	0.497 ± 0.001	1511.636 ± 238.852	18.591 ± 7.582
$w = 4, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	18.916 ± 2.105	0.929 ± 0.015	0.065 ± 0.014	0.497 ± 0.002	1533.566 ± 241.442	18.524 ± 7.587
$w = 5, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	18.885 ± 2.121	0.931 ± 0.014	0.064 ± 0.014	0.498 ± 0.002	1541.592 ± 232.272	17.907 ± 7.508
$w = 6, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.269 ± 2.215	0.933 ± 0.014	0.063 ± 0.014	0.498 ± 0.002	1560.129 ± 235.673	18.258 ± 7.769
$w = 7, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.213 ± 2.216	0.933 ± 0.014	0.063 ± 0.014	0.498 ± 0.002	1441.079 ± 200.736	18.294 ± 7.762
$w = 8, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.409 ± 2.254	0.934 ± 0.014	0.062 ± 0.014	0.498 ± 0.002	1414.262 ± 184.353	18.265 ± 7.730
$w = 9, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.504 ± 2.266	0.936 ± 0.013	0.062 ± 0.014	0.499 ± 0.002	1472.411 ± 190.027	18.194 ± 7.701
$w = 10, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.647 ± 2.317	0.937 ± 0.013	0.061 ± 0.014	0.499 ± 0.002	1488.688 ± 196.648	18.371 ± 7.947
$w = 11, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.862 ± 2.358	0.936 ± 0.013	0.061 ± 0.014	0.498 ± 0.002	1403.129 ± 169.845	18.420 ± 7.916
$w = 12, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.934 ± 2.373	0.934 ± 0.014	0.061 ± 0.014	0.497 ± 0.002	1444.162 ± 174.673	18.907 ± 8.372
$w = 13, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.957 ± 2.368	0.936 ± 0.013	0.062 ± 0.014	0.499 ± 0.002	1465.005 ± 187.948	19.086 ± 8.590
$w = 14, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.066 ± 2.423	0.937 ± 0.013	0.061 ± 0.014	0.499 ± 0.002	1514.741 ± 188.504	19.285 ± 8.865
$w = 15, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.195 ± 2.426	0.939 ± 0.013	0.061 ± 0.014	0.500 ± 0.002	1443.138 ± 183.408	19.234 ± 8.805
$w = 16, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.253 ± 2.458	0.939 ± 0.013	0.061 ± 0.014	0.500 ± 0.002	1500.906 ± 199.944	19.259 ± 8.831
$w = 17, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.335 ± 2.492	0.941 ± 0.013	0.061 ± 0.014	0.501 ± 0.001	1578.627 ± 215.405	19.258 ± 8.951
$w = 18, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.049 ± 2.342	0.937 ± 0.014	0.060 ± 0.013	0.499 ± 0.002	1639.302 ± 231.755	18.904 ± 8.652
$w = 19, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.021 ± 2.339	0.936 ± 0.013	0.059 ± 0.013	0.498 ± 0.001	1594.704 ± 213.182	19.030 ± 8.698
$w = 20, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.220 ± 2.374	0.938 ± 0.012	0.059 ± 0.013	0.499 ± 0.001	1584.171 ± 227.209	18.880 ± 8.602
$w = 1, \mathbf{v}$	9.118 ± 1.320	0.926 ± 0.015	0.065 ± 0.014	0.496 ± 0.002	1649.606 ± 262.578	0.255 ± 7.292
$w = 2, \mathbf{v}$	9.063 ± 1.464	0.923 ± 0.015	0.064 ± 0.014	0.494 ± 0.002	1587.528 ± 251.271	0.412 ± 7.491
$w = 3, \mathbf{v}$	9.620 ± 1.385	0.925 ± 0.015	0.063 ± 0.013	0.494 ± 0.002	1552.462 ± 239.063	0.531 ± 7.585
$w = 4, \mathbf{v}$	9.288 ± 1.260	0.923 ± 0.015	0.065 ± 0.014	0.494 ± 0.002	1593.134 ± 243.301	0.470 ± 7.588
$w = 5, \mathbf{v}$	9.901 ± 1.424	0.926 ± 0.015	0.063 ± 0.014	0.495 ± 0.001	1596.220 ± 234.643	0.859 ± 7.508
$w = 6, \mathbf{v}$	10.761 ± 1.867	0.928 ± 0.015	0.063 ± 0.014	0.495 ± 0.001	1609.972 ± 236.479	18.216 ± 7.768
$w = 7, \mathbf{v}$	11.300 ± 2.067	0.927 ± 0.015	0.062 ± 0.014	0.495 ± 0.001	1514.797 ± 207.056	18.247 ± 7.765
$w = 8, \mathbf{v}$	11.502 ± 2.181	0.929 ± 0.015	0.062 ± 0.014	0.496 ± 0.001	1495.852 ± 190.094	18.224 ± 7.730
$w = 9, \mathbf{v}$	12.297 ± 2.135	0.932 ± 0.014	0.062 ± 0.014	0.497 ± 0.001	1557.430 ± 201.636	18.148 ± 7.704
$w = 10, \mathbf{v}$	13.405 ± 2.485	0.934 ± 0.014	0.062 ± 0.014	0.498 ± 0.001	1554.391 ± 196.343	18.341 ± 7.949
$w = 11, \mathbf{v}$	12.810 ± 2.724	0.935 ± 0.014	0.061 ± 0.014	0.498 ± 0.001	1496.342 ± 175.994	18.380 ± 7.919
$w = 12, \mathbf{v}$	13.751 ± 2.810	0.930 ± 0.015	0.061 ± 0.014	0.496 ± 0.001	1516.673 ± 181.557	18.872 ± 8.375
$w = 13, \mathbf{v}$	14.991 ± 3.052	0.930 ± 0.014	0.062 ± 0.014	0.496 ± 0.001	1513.243 ± 188.059	19.051 ± 8.592
$w = 14, \mathbf{v}$	15.472 ± 3.232	0.931 ± 0.015	0.061 ± 0.014	0.496 ± 0.002	1588.332 ± 192.987	19.247 ± 8.867
$w = 15, \mathbf{v}$	14.162 ± 2.966	0.933 ± 0.015	0.061 ± 0.014	0.497 ± 0.001	1500.391 ± 185.857	19.197 ± 8.808
$w = 16, \mathbf{v}$	15.280 ± 3.294	0.932 ± 0.015	0.061 ± 0.014	0.496 ± 0.002	1532.484 ± 197.135	19.234 ± 8.834
$w = 17, \mathbf{v}$	14.257 ± 3.023	0.934 ± 0.015	0.061 ± 0.014	0.497 ± 0.002	1638.960 ± 218.039	19.235 ± 8.953

	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Visit/Post
$w = 18, \mathbf{v}$	14.263 ± 2.686	0.932 ± 0.015	0.060 ± 0.014	0.496 ± 0.002	1700.408 ± 235.241	18.876 ± 8.655
$w = 19, \mathbf{v}$	15.359 ± 3.037	0.929 ± 0.014	0.059 ± 0.013	0.494 ± 0.002	1648.964 ± 215.476	19.005 ± 8.702
$w = 20, \mathbf{v}$	16.942 ± 3.160	0.929 ± 0.014	0.059 ± 0.013	0.494 ± 0.002	1653.162 ± 230.106	18.859 ± 8.606
$w = 1, \mathbf{v}, \mathbf{t}_\triangle$	14.141 ± 1.764	0.930 ± 0.014	0.066 ± 0.014	0.498 ± 0.002	1537.992 ± 251.250	18.272 ± 7.291
$w = 2, \mathbf{v}, \mathbf{t}_\triangle$	18.211 ± 2.102	0.928 ± 0.014	0.066 ± 0.014	0.497 ± 0.002	1550.576 ± 251.645	18.463 ± 7.491
$w = 3, \mathbf{v}, \mathbf{t}_\triangle$	18.740 ± 2.127	0.929 ± 0.014	0.065 ± 0.014	0.497 ± 0.002	1511.829 ± 238.838	18.584 ± 7.582
$w = 4, \mathbf{v}, \mathbf{t}_\triangle$	18.906 ± 2.105	0.929 ± 0.015	0.065 ± 0.014	0.497 ± 0.002	1533.555 ± 241.443	18.524 ± 7.587
$w = 5, \mathbf{v}, \mathbf{t}_\triangle$	18.874 ± 2.121	0.931 ± 0.014	0.064 ± 0.014	0.498 ± 0.002	1541.587 ± 232.271	17.907 ± 7.508
$w = 6, \mathbf{v}, \mathbf{t}_\triangle$	19.250 ± 2.213	0.933 ± 0.014	0.063 ± 0.014	0.498 ± 0.002	1560.129 ± 235.673	18.258 ± 7.769
$w = 7, \mathbf{v}, \mathbf{t}_\triangle$	19.203 ± 2.216	0.933 ± 0.014	0.063 ± 0.014	0.498 ± 0.002	1441.079 ± 200.736	18.294 ± 7.762
$w = 8, \mathbf{v}, \mathbf{t}_\triangle$	19.396 ± 2.253	0.934 ± 0.014	0.062 ± 0.014	0.498 ± 0.002	1414.262 ± 184.353	18.265 ± 7.730
$w = 9, \mathbf{v}, \mathbf{t}_\triangle$	19.501 ± 2.265	0.936 ± 0.013	0.062 ± 0.014	0.499 ± 0.002	1472.411 ± 190.027	18.194 ± 7.701
$w = 10, \mathbf{v}, \mathbf{t}_\triangle$	19.642 ± 2.316	0.937 ± 0.013	0.061 ± 0.014	0.499 ± 0.002	1488.669 ± 196.646	18.371 ± 7.947
$w = 11, \mathbf{v}, \mathbf{t}_\triangle$	19.862 ± 2.358	0.936 ± 0.013	0.061 ± 0.014	0.498 ± 0.002	1403.129 ± 169.845	18.420 ± 7.916
$w = 12, \mathbf{v}, \mathbf{t}_\triangle$	19.933 ± 2.373	0.934 ± 0.014	0.061 ± 0.014	0.497 ± 0.002	1444.157 ± 174.672	18.907 ± 8.372
$w = 13, \mathbf{v}, \mathbf{t}_\triangle$	19.953 ± 2.367	0.936 ± 0.013	0.062 ± 0.014	0.499 ± 0.002	1464.990 ± 187.946	19.086 ± 8.590
$w = 14, \mathbf{v}, \mathbf{t}_\triangle$	20.062 ± 2.423	0.937 ± 0.013	0.061 ± 0.014	0.499 ± 0.002	1514.729 ± 188.503	19.285 ± 8.865
$w = 15, \mathbf{v}, \mathbf{t}_\triangle$	20.193 ± 2.426	0.939 ± 0.013	0.061 ± 0.014	0.500 ± 0.002	1443.130 ± 183.407	19.234 ± 8.805
$w = 16, \mathbf{v}, \mathbf{t}_\triangle$	20.250 ± 2.458	0.939 ± 0.013	0.061 ± 0.014	0.500 ± 0.002	1500.898 ± 199.942	19.259 ± 8.831
$w = 17, \mathbf{v}, \mathbf{t}_\triangle$	20.334 ± 2.492	0.941 ± 0.013	0.061 ± 0.014	0.501 ± 0.001	1578.624 ± 215.404	19.258 ± 8.951
$w = 18, \mathbf{v}, \mathbf{t}_\triangle$	20.048 ± 2.342	0.937 ± 0.014	0.060 ± 0.013	0.499 ± 0.002	1639.300 ± 231.754	18.904 ± 8.652
$w = 19, \mathbf{v}, \mathbf{t}_\triangle$	20.021 ± 2.339	0.936 ± 0.013	0.059 ± 0.013	0.498 ± 0.001	1594.704 ± 213.182	19.030 ± 8.698
$w = 20, \mathbf{v}, \mathbf{t}_\triangle$	20.220 ± 2.374	0.938 ± 0.012	0.059 ± 0.013	0.499 ± 0.001	1584.171 ± 227.209	18.880 ± 8.602

Overall average and window average perform significantly worst than the learnt models, as reflected in both the MAPE and the T-score. There is also a slight improvement in the Pr_{error} in the learnt models.

Taking into account the T-score and the number of visits together, would seem that $\mathbf{t}_\triangle$, features representing the previous time intervals, are important features when determining the next time interval. In the absence of these features, we observe that the T-score increases by about 1%. In this experiment, we use purely word frequency features. This gives only a slight improvent over not using them.

High values for Pr_{miss} and low for Pr_{fa}, are due to Pr_{miss} being conditioned on there being a post within the window. Since the posts come in bursts, visits are fairly periodic, and intervals between visits are larger than post bursts. When there are more posts than visits in windows with posts, we have higher Pr_{miss}

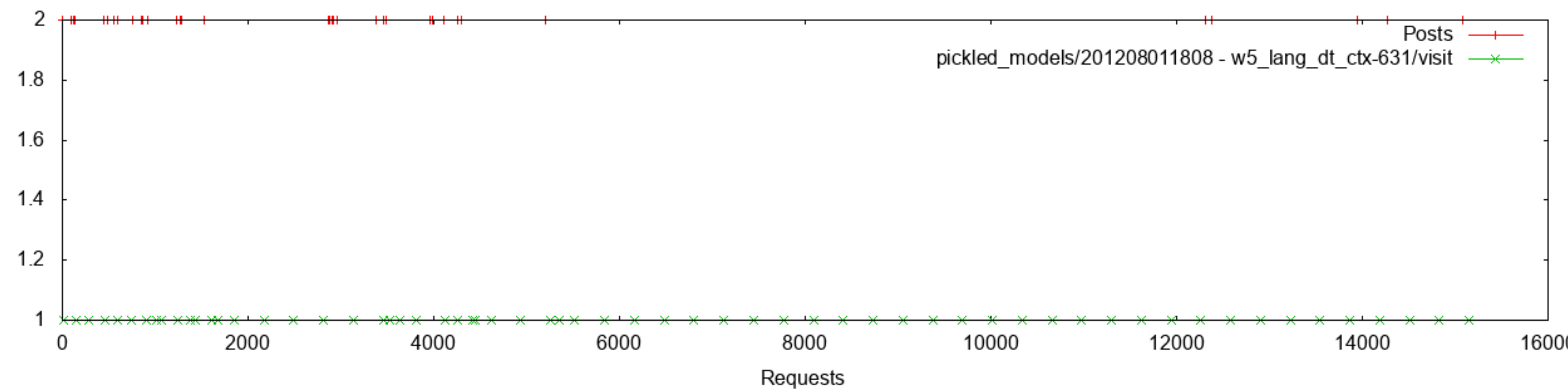

Figure 2: Visitation chart for a model using the $w = 5, \mathbf{t}_\Delta, \mathbf{t}_{\mathrm{ctx}}, \mathbf{w}$ feature set. Invalid Predictions = 0.758, $Pr_{error} = 0.485$, T-score = 119.612, Posts = 41, Visits = 62

2.1 Discounted sum of previous instances

The current method uses only information on the current w posts. However, posts made further in the history of the thread may have an effect on when the latest posts arrive. The magnitude of this effect, however, may diminish over time.

Following this intuition we attempt to use a discounted sum over previous posts' word frequency vector:

$$\mathbf{x}_t = \mathbf{v}_t + \alpha \mathbf{x}_{t-1}$$

where $\mathbf{x}_t$ is the feature vector at post t, and $\mathbf{v}$ is the word frequency vector. α is the *discount factor* and satisfies $0 \leq \alpha < 1$.

2.2 Stochastic Gradient Descent

We attempt to use stochastic gradient descent to estimate the function f. However, during runtime, instead of using a static function, we continue to allow f to vary whenever new posts and their update times are observed. Since $f(\mathbf{x}_{t-w}, \ldots, \mathbf{x}_{t-1}) > 0$, we used a scaled sigmoid function,

$$f(\mathbf{X}) = \frac{\Lambda - \lambda}{1 + e^{\mathbf{w} \cdot \mathbf{X}}} + \lambda$$

where Λ and λ are the scaling factors. This results in $f : \mathbb{R}^{|\mathbf{X}|} \to (\lambda, \Lambda)$. Bounding the estimation function between λ and Λ allows us to restrict the prediction from becoming negative, or, becoming exceedingly huge. For our purposes, we set $\lambda = Q_3 + 2.5(Q_3 - Q_1)$, where Q_n is the value at the n-th quartile.

The resulting update rule for $\mathbf{w}$ is then given by,

$$\Delta \mathbf{w}_i = \eta \underbrace{\left(\widehat{\Delta_t} - \Delta_t\right)}_{\text{error term}} \underbrace{(f(\mathbf{X})(1 - f(\mathbf{X})))}_{\text{gradient}} \mathbf{X}_i$$

which is similar to the delta update rule found in artificial neural networks. We omit

	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Visit/Post
Average $w = 5$	330.285	0.951	0.054	0.502	6418.208	16.464
Average $w = 10$	305.557	0.955	0.053	0.504	4598.955	17.291
Average $w = 15$	308.547	0.954	0.054	0.504	3833.605	18.337
Average $w = 20$	265.124	0.953	0.054	0.504	3340.929	18.102
Average $w = 25$	257.844	0.955	0.052	0.503	3186.309	17.927
Average $w = 30$	244.988	0.957	0.050	0.504	2859.380	18.362
$w = 5, \mathbf{t}_\Delta$	18.884	0.931	0.064	0.498	1541.595	17.907
$w = 10, \mathbf{t}_\Delta$	19.647	0.937	0.061	0.499	1488.688	18.371
$w = 15, \mathbf{t}_\Delta$	20.195	0.939	0.061	0.500	1443.138	19.234
$w = 20, \mathbf{t}_\Delta$	20.220	0.938	0.059	0.499	1584.171	18.880
$w = 25, \mathbf{t}_\Delta$	20.953	0.937	0.056	0.496	1649.098	18.612
$w = 30, \mathbf{t}_\Delta$	21.242	0.941	0.054	0.498	1626.782	18.984
$w = 5, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	18.885	0.931	0.064	0.498	1541.592	17.907
$w = 10, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	19.647	0.937	0.061	0.499	1488.688	18.371
$w = 15, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.195	0.939	0.061	0.500	1443.138	19.234
$w = 20, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.220	0.938	0.059	0.499	1584.171	18.880
$w = 25, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	20.953	0.937	0.056	0.496	1649.098	18.612
$w = 30, \mathbf{t}_\Delta, \mathbf{t}_{\text{ctx}}$	21.242	0.941	0.054	0.498	1626.782	18.984
$w = 5, \mathbf{v}$	9.382	0.923	0.063	0.493	1597.533	17.862
$w = 10, \mathbf{v}$	13.863	0.934	0.061	0.498	1551.375	18.339
$w = 15, \mathbf{v}$	13.217	0.934	0.060	0.497	1507.589	19.182
$w = 20, \mathbf{v}$	14.849	0.930	0.059	0.494	1630.643	18.848
$w = 25, \mathbf{v}$	17.542	0.930	0.055	0.493	1700.990	18.579
$w = 30, \mathbf{v}$	18.627	0.937	0.054	0.496	1653.156	18.959
$w = 5, \mathbf{v}, \mathbf{t}_\Delta$	18.877	0.931	0.064	0.498	1541.588	17.907
$w = 10, \mathbf{v}, \mathbf{t}_\Delta$	19.645	0.937	0.061	0.499	1488.680	18.371
$w = 15, \mathbf{v}, \mathbf{t}_\Delta$	20.193	0.939	0.061	0.500	1443.130	19.234
$w = 20, \mathbf{v}, \mathbf{t}_\Delta$	20.220	0.938	0.059	0.499	1584.171	18.880
$w = 25, \mathbf{v}, \mathbf{t}_\Delta$	20.953	0.937	0.056	0.496	1649.098	18.612
$w = 30, \mathbf{v}, \mathbf{t}_\Delta$	21.242	0.941	0.054	0.498	1626.782	18.984

Table 4: Experiment results: Varying feature sizes

	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Visit/Post
Average $w = 5$	330.285	0.974	0.010	0.492	23941.231	0.620
Average $w = 10$	305.557	0.976	0.010	0.493	24932.464	0.635
Average $w = 15$	308.547	0.975	0.010	0.493	27999.236	0.639
Average $w = 20$	265.124	0.975	0.010	0.493	25785.328	0.651
Average $w = 25$	257.844	0.977	0.010	0.494	27307.482	0.658
Average $w = 30$	244.988	0.979	0.010	0.495	26039.058	0.655
$w = 5, \mathbf{t}_\Delta$	18.884	0.975	0.011	0.493	23348.003	0.697
$w = 10, \mathbf{t}_\Delta$	19.647	0.974	0.011	0.492	24208.018	0.704
$w = 15, \mathbf{t}_\Delta$	20.195	0.976	0.011	0.494	27456.946	0.706
$w = 20, \mathbf{t}_\Delta$	20.220	0.976	0.011	0.494	25246.673	0.720
$w = 25, \mathbf{t}_\Delta$	20.953	0.975	0.011	0.493	26976.438	0.719
$w = 30, \mathbf{t}_\Delta$	21.242	0.975	0.011	0.493	25776.611	0.724
$w = 5, \mathbf{t}_\Delta, \mathbf{t}_{\mathrm{ctx}}$	18.885	0.975	0.011	0.493	23348.002	0.697
$w = 10, \mathbf{t}_\Delta, \mathbf{t}_{\mathrm{ctx}}$	19.647	0.974	0.011	0.492	24208.018	0.704
$w = 15, \mathbf{t}_\Delta, \mathbf{t}_{\mathrm{ctx}}$	20.195	0.976	0.011	0.494	27456.946	0.706
$w = 20, \mathbf{t}_\Delta, \mathbf{t}_{\mathrm{ctx}}$	20.220	0.976	0.011	0.494	25246.673	0.720
$w = 25, \mathbf{t}_\Delta, \mathbf{t}_{\mathrm{ctx}}$	20.953	0.975	0.011	0.493	26976.438	0.719
$w = 30, \mathbf{t}_\Delta, \mathbf{t}_{\mathrm{ctx}}$	21.242	0.975	0.011	0.493	25776.611	0.724
$w = 5, \mathbf{v}$	9.382	0.969	0.011	0.490	23380.715	0.686
$w = 10, \mathbf{v}$	13.863	0.971	0.011	0.491	24258.557	0.698
$w = 15, \mathbf{v}$	13.217	0.973	0.010	0.492	27518.222	0.689
$w = 20, \mathbf{v}$	14.849	0.970	0.011	0.490	25276.263	0.711
$w = 25, \mathbf{v}$	17.542	0.970	0.011	0.490	27016.238	0.710
$w = 30, \mathbf{v}$	18.627	0.972	0.011	0.491	25792.662	0.714
$w = 5, \mathbf{v}, \mathbf{t}_\Delta$	18.877	0.975	0.011	0.493	23347.998	0.697
$w = 10, \mathbf{v}, \mathbf{t}_\Delta$	19.645	0.974	0.011	0.492	24208.011	0.704
$w = 15, \mathbf{v}, \mathbf{t}_\Delta$	20.193	0.976	0.011	0.494	27456.937	0.706
$w = 20, \mathbf{v}, \mathbf{t}_\Delta$	20.220	0.976	0.011	0.494	25246.673	0.720
$w = 25, \mathbf{v}, \mathbf{t}_\Delta$	20.953	0.975	0.011	0.493	26976.438	0.719
$w = 30, \mathbf{v}, \mathbf{t}_\Delta$	21.242	0.975	0.011	0.493	25776.611	0.724

Table 5: Experiment results: Using exponential increase

	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Posts	Visits
$\alpha = 0.0, \mathbf{v}$	13.217	0.934	0.060	0.497	1507.589	30.402	528.247
$\alpha = 0.1, \mathbf{v}$	13.626	0.932	0.060	0.496	1500.257	30.402	529.608
$\alpha = 0.2, \mathbf{v}$	14.021	0.934	0.060	0.497	1487.145	30.402	530.175
$\alpha = 0.3, \mathbf{v}$	14.446	0.933	0.060	0.497	1483.811	30.402	530.825
$\alpha = 0.4, \mathbf{v}$	14.612	0.933	0.060	0.497	1449.927	30.402	531.742
$\alpha = 0.5, \mathbf{v}$	14.579	0.931	0.061	0.496	1434.925	30.402	532.701
$\alpha = 0.6, \mathbf{v}$	14.668	0.934	0.061	0.497	1445.370	30.402	533.165
$\alpha = 0.7, \mathbf{v}$	14.907	0.931	0.061	0.496	1442.314	30.402	534.629
$\alpha = 0.8, \mathbf{v}$	16.208	0.933	0.060	0.497	1440.828	30.402	532.680
$\alpha = 0.9, \mathbf{v}$	18.595	0.931	0.061	0.496	1472.814	30.402	532.691

Table 6: Experiment results: Window size of 15, using discounted sum feature vector at $t - 1$.

	MAPE	Pr_{miss}	Pr_{fa}	Pr_{error}	T-score	Visit/Post
$\eta = 5 \cdot 10^{-1}, \mathbf{v}$	28.208	0.929	0.082	0.506	1025.388	27.703
$\eta = 5 \cdot 10^{-2}, \mathbf{v}$	36.432	0.943	0.064	0.504	1293.619	21.060
$\eta = 5 \cdot 10^{-3}, \mathbf{v}$	54.939	0.945	0.059	0.502	1371.346	19.422
$\eta = 5 \cdot 10^{-4}, \mathbf{v}$	72.015	0.942	0.059	0.501	1475.683	19.153
$\eta = 5 \cdot 10^{-5}, \mathbf{v}$	73.763	0.940	0.059	0.499	1595.563	19.097
$\eta = 5 \cdot 10^{-6}, \mathbf{v}$	77.506	0.942	0.059	0.501	1525.705	19.122
$\eta = 5 \cdot 10^{-7}, \mathbf{v}$	82.477	0.944	0.059	0.502	1440.440	19.121
$\eta = 5 \cdot 10^{-8}, \mathbf{v}$	87.682	0.944	0.059	0.501	1407.172	19.108
$\eta = 5 \cdot 10^{-9}, \mathbf{v}$	93.382	0.944	0.059	0.502	1416.182	19.110
$\eta = 5 \cdot 10^{-10}, \mathbf{v}$	101.881	0.944	0.059	0.501	1451.729	19.106
$\eta = 5 \cdot 10^{-11}, \mathbf{v}$	107.706	0.943	0.059	0.501	1482.868	19.104
$\eta = 5 \cdot 10^{-12}, \mathbf{v}$	108.455	0.944	0.059	0.501	1487.555	19.104

Table 7: Experiment results: Using stochastic gradient descent, with $\lambda = 0$

the scaling factor in the gradient as it is a constant and then experiment with various values of η, the learning rate. The results are seen in Table 7