Twitter-based Mood Evaluation

Vishnu Sresht
Email: vishnusr@mit.edu

Pratik Chaudhari
Email: pratikac@mit.edu

Abstract—By providing a convenient, readily-accessible way to reach out to a global audience, twitter has revolutionized the way we broadcast our feelings to the rest of the world. In addition to its empowerment of the diva within us all, Twitter's meteoric rise in popularity has enabled every one of us to listen to the voices of millions and comprehend mankind's *zeitgest* at an uprecedented resolution. However, the deluge of verbiage unleashed by Twitter's ubiquitous usage must be tamed before its wealth of information can be exploited for sociologically-beneficial research. In this project, we attempt to survey, compare and contrast machine learning techniques for one particular form of large-scale tweet analysis - that of determining 'how positive people feel' at any given moment through the sentiment analysis of their tweets. To this end, we also introduce a novel source of already classified text corpora for use as training data. Sites like My Life Is G and FML are a hiherto unutilized source of crowd-curated, well classified training data of text snippets of appropriate length. In this paper, we intend to study the relative efficacies of several commonly employed classification algorithms, including Naive Bayes, Support Vector Machines, and K-Nearest Neighbor classifiers when applied to the task of of detecting the mood of the nation on a real-time basis.

I. INTRODUCTION

Kramer's theory provides a precise mathematical formulation and solution to a problem that recurs in various forms in several branches of the physical sciences: that of determining the kinetics of stochastically motivated barrier crossing. The simplest version of this problem deals with determing the average amount of time needed by a particle to surmount an energy barrier through thermal agitation.

Kramer's theory provides a precise mathematical formulation and solution to a problem that recurs in various forms in several branches of the physical sciences: that of determining the kinetics of stochastically motivated barrier crossing. The simplest version of this problem deals with determing the average amount of time needed by a particle to surmount an energy barrier through thermal agitation.

Given pairs of MoS_2 sheets with inter-sheet interaction energy $\Phi(z)$, we would like to determine the rate at which pairs transition from the metastable dispersed configuration to the stable agglomerated configuration.

In the limit of high viscous drag, the relative stochastic motion of two MoS_2 plates under the influence of the field $\Phi(z)$ is described by the following PDE:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = -\mathcal{D}\frac{\partial}{\partial z}\beta\Phi(z) + \sqrt{2\mathcal{D}}\xi(t) \tag{1}$$

where $\mathcal{D}$ is the diffusivity of the sheets in the surrounding medium, β is the inverse temperature $(k_B T)^{-1}$, and $\xi(t)$ corresponds to stochastic disturbances arising out of brownian motion and satisfies $\langle \xi(t)\xi(t')\rangle = \delta(t-t')$.

This PDE in terms of $z(t)$ can be written in terms of the probability $P(z,t)$ of finding the sheets at a separation z, at time t:

$$\frac{\partial}{\partial t}P(z) = \mathcal{D}\frac{\partial}{\partial z}\left[\frac{\partial}{\partial z}\beta\Phi(z) + \frac{\partial}{\partial z}\right]P(z) \tag{2}$$

We could write Newton's second law for the pair of plates

$$
\begin{aligned}
M\ddot{z} &= F_{\text{total}}\\
&= F_{\text{PMF}} + F_{\text{viscous}} + F_{\text{brownian}}\\
\Rightarrow M\ddot{z} &= -\frac{\partial}{\partial z}\phi(z) - \zeta\frac{\mathrm{d}x}{\mathrm{d}t} + \xi(t)
\end{aligned} \tag{3}
$$

Note that the term F_{viscous} corresponds to the viscous resistance or linear drag that is experienced by bodies moving slowly through viscous media where the Reynolds number, $R_e < 1$. This leads us to the Langevin equations for the dynamics of the system in terms of the intersheet separation z:

$$
\begin{aligned}
\dot{z} &= v\\
\dot{v} &= -\frac{1}{M}\frac{\partial}{\partial z}\phi(z) - \frac{\zeta}{M}v + \frac{1}{M}\xi(t)
\end{aligned} \tag{4}
$$

Over the time scale of the aggregation process, we expect the MoS_2 sheets to sample all possible relative orientations. Consequently, these sheets can be modeled as spheres with an equivalent hydrodynamic radius of $R_d = \sqrt{A/\pi}$.