

TRUNCATED GAUSSIAN PROCESS REGRESSION FOR PREDICTING GROWTH
OF ABDOMINAL AORTIC ANEURYSM AND FOR TEMPORAL MODELING OF
SENTIMENTS

By

Ahsan Ijaz

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical Engineering - Master of Science

2013

ABSTRACT

TRUNCATED GAUSSIAN PROCESS REGRESSION FOR PREDICTING GROWTH OF ABDOMINAL AORTIC ANEURYSM AND FOR TEMPORAL MODELING OF SENTIMENTS

By

Ahsan Ijaz

An abdominal Aortic Aneurysm (AAA) is a form of vascular disease causing focal enlargement of the abdominal aorta. As part of the present study, we use series of computer tomography scans (CT-scans) of small AAAs taken at different times to model and predict the spatio-temporal evolution of AAAs. Using the proposed methodology and available CT scan data, the prediction of an AAA can be made for any time using truncated Gaussian process regression. The results of our case study show excellent outcomes of our algorithms when they are compared to the true CT scan images.

Second part of the thesis concerns the temporal modeling of sentiments expressed through textual information in Social networks. As part of this study, we explore the issues related to the temporal models and provide an efficient method which overcomes the inefficiencies associated with traditional schemes. A nonparametric, computationally efficient temporal model is provided using truncated Gaussian process regression. The model is built so that a noise parameter is estimated using the sentiment classification error metrics and inserted in the regression setting. This makes the method generic and any form of quantification of sentiments (through manual labeling or by some other classification scheme) can be used with improvement on final results. Baseline sentiment analysis schemes are used in conjunction with the proposed temporal model on data crawled from Twitter to express the utility of the scheme.

TABLE OF CONTENTS

LIST OF TABLES	v
LIST OF FIGURES	vi
LIST OF ALGORITHMS	viii

Chapter 1 Prediction of Abdominal Aortic Aneurysms using Gaussian Process Regression	1
1.1 Introduction	1
1.1.1 Notation	5
1.1.2 Organization	6
1.2 Data and Methods	6
1.2.1 Data	6
1.2.2 Observations of Center Lines from the Data	7
1.2.3 Observations of AAA Surfaces from the Data	8
1.2.4 Gaussian Process Regression	10
1.3 Spatio-temporal Modeling of an AAA	12
1.3.1 Spatio-Temporal Modeling of the Center Line	13
1.3.2 AAA Surface Prediction	15
1.4 Case Study and Results	17
1.5 Discussion	19
1.5.1 Decision Making via Prediction and its Confidence Region	19
1.5.2 Scheduling of CT Scans	20
1.5.3 Hyperparameters as Possible Feature Vectors	21
1.5.4 Limitations and Future Research Directions	21
1.6 Conclusion	23

Chapter 2 Temporal Modeling and Forecasting of Sentiments in Online Social Networks	31
2.1 Related Work	33
2.1.1 Sentiment Analysis	34
2.1.2 Temporal Modeling	34
2.1.3 Notation	35
2.2 Proposed Method	36
2.2.1 Sentiment Classification	36
2.2.2 Error Characterization	37
2.2.2.1 Sampling and Scalability	40
2.2.2.2 Temporal Model	41
2.3 Experimental setup and Data	41

2.3.1	Training Data and Feature Extraction	42
2.4	Results	44
2.4.1	Classification Results	44
2.4.2	Gaussian process based Temporal Model	45
2.4.3	Effects of Sampling	48
2.5	Conclusion and Future Work	50
APPENDICES		51
	Appendix A: Generation of samples of the center line	52
	Appendix B: Surface parameterization	55
BIBLIOGRAPHY		58

LIST OF TABLES

Table 1.1	Scan times for Patient given as days after first scan	6
Table 1.2	Hyperparameters For Surface Prediction	16
Table 1.3	Error measures in Prediction using data of Patient B	18
Table 2.1	Twitter data collected	42
Table 2.2	Training Data using emoticons	43
Table 2.3	Classification Results	45
Table 2.4	Most informative features	46

LIST OF FIGURES

Figure 1.1	Parametrized axis system $r(s, \theta)$ (black) and center line (red), where s is the travel length in mm along the center line and $\theta \in \Theta$ is the angle in radians. The output of r is given as the distance from the center line to the point on the surface. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.	9
Figure 1.2	The predicted center line for fourth scan using first three scans. The predicted center line is shown in green and the original center line is in blue.	14
Figure 1.3	Case 1: (a) Parameterized surface using original data for time t_3 . (b) Parameterized surface using results of prediction for time t_3	24
Figure 1.4	Case 1: (a) Original surface of aorta at time t_3 . (b) Reconstructed image of aorta using predicted surface and center line for time t_3	25
Figure 1.5	Case 2: (a) Parameterized surface using original data for time t_4 . (b) Parameterized surface using results of prediction for time t_4	26
Figure 1.6	Case 2: (a) Original surface of aorta at time t_4 . (b) Reconstructed image of aorta using predicted surface and center line for time t_4	27
Figure 1.7	Case 3: (a) Parameterized surface using original data for time t_3 . (b) Parameterized surface using results of interpolation for time t_3	28
Figure 1.8	Case 3: (a) Original surface of aorta at time t_3 . (b) Reconstructed image of aorta using interpolated surface and center line for time t_3	29
Figure 1.9	Case 2: Predicted surface (middle) with confidence intervals (up and down) at time t_4	30
Figure 2.1	Aggregated sentiments for Obama(blue) and Romney(red)	45
Figure 2.2	Predicted Target Sentiments(Green) with Aggregated sentiment(Blue) for Obama	47

Figure 2.3	Predicted function(green) for Obama sentiments(blue) with confidence interval(grey)	48
Figure 2.4	Predicted function(green) for Romney sentiments(red) with confidence interval(grey)	49
Figure 2.5	Decrease of mean square error from Original temporal sentiments with increase of Samples	49

LIST OF ALGORITHMS

Algorithm 1	Generation of Samples of Center Line	54
Algorithm 2	Surface Parameterization	56

Chapter 1

Prediction of Abdominal Aortic Aneurysms using Gaussian Process Regression

1.1 Introduction

The aorta is a major artery in which blood circulates through the heart. An aortic aneurysm is identified as enlargement of the aorta greater than 50% of the normal diameter. The vast majority of aortic aneurysms are in the abdominal region (AAAs) and over 90% of these AAAs occur specifically within the infrarenal aorta [1] [2] where a diameter greater than 3 cm is considered as an aneurysm. The infrarenal aorta is the section of the abdominal aorta which lies between the renal branches and the iliac bifurcation. AAAs are a serious medical condition that, when left untreated, can cause vessel rupture with patient mortality rates more than 95% [3] [4]. Therefore, a thorough understanding of expansion and rupture of AAAs is desired.

The expansion rate and rupture potential of AAAs have been associated by several studies. Multifaceted biological processes have been identified as affecting the growth of AAAs including biochemical, biomechanical, cellular, and proteolytic factors [5]. Morphological

features and regional geometrical variations have been analyzed. Relations between wall stress distribution and hemodynamic factors have been studied using fluid structure interaction (FSI) for ruptured aneurysms [6], in presence of intraluminal thrombus (ILT) and as a result of morphology and blood pressure [7–9]. Findings over recent decades have also demonstrated that the vascular tissues exhibit a remarkable ability of adaptation under various physiological conditions [10–12]. In particular, blood vessels seek to maintain a preferred (homeostatic) mechanical state in conditions of altered blood flow [13, 14], blood pressure [15, 16], axial extension [17–19], and during disease processes. Based on the increased understanding of vascular diseases and advances in theoretical and computational modeling of vascular adaptation, many computational models have been developed to describe vascular adaptation in various physiological and pathological conditions such as altered blood flow, sustained axial stretch, hypertension, intracranial aneurysm, and vasospasm [20–23]. The earlier studies, however, have focused mainly on hypothesis-testing, i.e., proving the feasibility of stress-mediated mechanisms in vascular adaptation that have been proposed as hypotheses, mostly with simple geometries.

Some recent works have been conducted on growth and remodeling (G&R) of AAAs based on patient-specific geometry as well. Zeinali-Davarani *et al.* [24] developed the G&R model to account for both elastic degeneration and stress-mediated collagen turnover during AAA development using finite element analysis (FEA). A coupled simulation of G&R with hemodynamics was conducted for studying its effects on AAA expansion [25]. Geometric, kinetic and material parameters have also been identified for individual patients using inverse optimization techniques for modeling the growth of AAAs. For the estimation of constitutive material parameters of the artery, a nonlinear estimation method has been suggested [26]. In addition, it is reported that kinetic parameters such as collagen turnover, rates of pro-

duction, half-life deposition stretch and material stiffness depends strongly spatio-temporal changes in wall thickness, biaxial stresses and maximum collagen stretch [27] [28]. Spatial distribution of the thickness and material properties of porcine thoracic aortas have also been investigated experimentally by extension-inflation tests with a stereo-vision system [29]. Furthermore, it has been shown that the same material parameters for AAA expansion can predict the intrasac-pressure dependent vascular adaptation after endovascular repair [30]. Transferring such advances in computational modeling of vascular diseases into an individualized predictive tool in clinical treatment, however, requires a major paradigm-shift due to the incompleteness of the model, limited information, and uncertainty associated with clinical measurement.

While the patient-specific computer models of G&R for an AAA provide insight into the associated risks, dependence on multitude of factors can obfuscate the prediction results. For a more precise feature independent prediction, we analyze the spatio-temporal patient-specific geometrical variations from a purely statistical framework.

The goal of this chapter is to provide a data-driven, nonparametric statistical framework using patient-specific data for improving the prediction results for aneurysm growth, hence assisting clinical management planning. While a computational G&R model is not utilized, the nonparametric modeling approach (using Gaussian process regression) in this study can be viewed as a step towards a Bayesian approach that will be capable of incorporating various uncertainties, patient-specific data, and computational models for G&R. To this end, for example, the computational G&R model can also be modeled in a nonparametric fashion.

In this chapter, the longitudinal patient-specific data used for the study consists of four CT scan images of an AAA. See more details about the data in Section 1.2.1. To achieve

our goal, first parameterization of data is developed as follows. Segmentation of an AAA is followed with image registration taking lumbar vertebrae as the static reference. The center line of an AAA for each scan is obtained and used for surface parameterization. The parameterization is carried out so that the distance of a point (on the AAA surface) from center line r normal to the center line is a function of length s , angle θ and time t of a scan as shown in Fig. 1.1. The surface prediction to assess the condition of an AAA for a future time is carried out using truncated spatio-temporal Gaussian process regression [31]. This needs to combine two predictions of the center line and the AAA surface. Using longitudinal data of AAA scans is desirable as it would reveal point wise progression of surface along with the associated uncertainties of prediction for any time of interest. Thus, both local and global changes are observed and the risk of rupture for small AAAs is also highlighted. The rate of progression is expressed in the model by hyperparameters of spatio-temporal Gaussian process which in turn are estimated using the maximum likelihood estimator [32]. Predicted AAAs in different cases are compared to existing true CT images to evaluate the performance of the proposed approach. A preliminary study about the technique using a different patient with two scans without the prediction of center lines, reconstruction of AAA surface, cases for validation and comparison between predicted images and original data was reported in [33].

In summary, the main contributions of this chapter are as follows.

- *Surface parameterization:* A unique surface parameterization for AAAs for visualization and analysis is developed.
- *Prediction of the center line of an AAA:* The temporal-variations in the center line of AAAs are used to develop a mathematical model to get a statistical estimation model of the center line at desired time.

- *Prediction of an AAA surface:* A statistical model for a parameterized AAA surface (with respect to a center line) is developed using computationally efficient truncated Gaussian process regression.
- *Prediction of an AAA and its validation:* Predicted AAAs are validated for three different cases. Each case has a training data set that is a subset of valuable longitudinal data of four CT scan AAA images of a patient. Comparison results of predictions with respect to the true (not-used) scan images are provided to evaluate the accuracy of the proposed scheme.
- *Prediction uncertainty:* The point-wise confidence interval associated with prediction is obtained for the predicted AAA surface. Error estimates using available data is also carried out.
- *Possible utility of the methodology:* Possible utility of the proposed method is discussed from helping decision making to feature extraction applications.

To the best of our knowledge, this is the first study that predicts the AAA growth using available (patient-specific) CT scan data in a statistical perspective allowing uncertainty quantification in the predicted AAA.

1.1.1 Notation

Standard notation will be used throughout this chapter. Let $\mathbb{R}, \mathbb{R}_{\geq 0}, \mathbb{R}_{>0}$, and $\mathbb{Z}$ denote, respectively, the sets of real, non-negative real, positive real, and integer numbers. I_n denotes the identity matrix of size n . For column vectors $v_a \in \mathbb{R}^a, v_b \in \mathbb{R}^b$, and $v_c \in \mathbb{R}^c$, $\text{col}(v_a, v_b, v_c) := [v_a \ v_b \ v_c] \in \mathbb{R}^{a+b+c}$ stacks all vectors to create one column vector, and $\|v_a\|$ denotes the Euclidean norm (or vector 2-norm) of v_a . $|A|$ denotes the determinant of a matrix $A \in \mathbb{R}^{n \times n}$. Let $\mathbb{E}(z)$ and $\text{Var}(z)$ denote, respectively, the expectation and the variance of

random vector z . A random vector $z \in \mathbb{R}^q$, which is distributed by a multivariate Gaussian distribution of a mean $\mu \in \mathbb{R}^q$ and a variance $\Sigma \in \mathbb{R}^{q \times q}$, is denoted by $z \sim \mathcal{N}(\mu, \Sigma)$. The first derivative operator on $h := \mathbb{R}^m \rightarrow \mathbb{R}$ with respect to vector $s \in \mathbb{R}^m$ is as follows.

$$\nabla h(s) = \frac{\partial h(s)}{\partial s} = \left(\frac{\partial h(s)}{\partial s_1}, \dots, \frac{\partial h(s)}{\partial s_m} \right).$$

1.1.2 Organization

This chapter is organized as follows. Section 1.2 explains our data and methods in detail. Sections 1.2.2 and 1.2.3 describe how we obtain observations from the data. Our main method, Gaussian process regression, is introduced in Section 1.2.4. Section 1.3 illustrates spatio-temporal modeling of AAAs using observations and Gaussian process regression methods. Successful results from our methodology are illustrated under three different cases in Section 1.4. Discussion and conclusion are followed in Sections 1.5 and 1.6, respectively.

1.2 Data and Methods

1.2.1 Data

Table 1.1: Scan times for Patient given as days after first scan

Scan Number	Time of Scan
Scan 1	$t_1 = 0$
Scan 2	$t_2 = 386$
Scan 3	$t_3 = 756$
Scan 4	$t_4 = 1120$

To evaluate our model, we used longitudinal data of four CT scan images of a male patient of age 54 years. The resolution of these CT scans are approximately 0.7 mm per

pixel. Details about the time of scans are provided in Table 1.1. This study was subject to Internal Review Board (IRB) approvals at both Michigan State University and Seoul National University Hospital. No patient consent was necessary since the data was collected for a retrospective study. Three dimensional (3D) models are reconstructed from CT scans using Mimics (Materialise, Leuven, Belgium) to get the longitudinal model set using semi-automatic segmentation. The longitudinal model set is further subjected to global image registration with respect to lumbar vertebrae, which is assumed to be relatively unchanging over time. This provides the spatial transform which maps the positions and orientations of AAAs with respect to the lumbar vertebrae. Image registration allows for an accurate investigation of the true spatial differences between scans at different times. The vertebra of the first scan is selected as the reference and the vertebra of second scan along with associated lumen and tissue models are aligned according to it. This registration is important for building the statistical growth model of the AAA since the spatial points of an AAA for all times should be aligned for building an accurate temporal evolution model. Thus image registration allows for the unique visualization of the true spatial differences of the AAA geometry and offers insight into the surface evolution of an AAA. The collection of (point cloud) data sets obtained from four scan images is denoted by $D_{scan} := \{D_1, \dots, D_4\}$ for further development.

1.2.2 Observations of Center Lines from the Data

The center line of an AAA acts as a reference for surface parameterization and analyzing morphological features. To obtain the center line, an iterative algorithm is developed for generating the center line for an arterial surface by collecting the center points of maximally inscribed spheres within the surface boundaries at fixed lengths. Using these center points of

spheres and 4^{th} order polynomial basis functions, a smooth line approximation of the center line is obtained as a function of length of an AAA. The algorithm is discussed in detail in Appendix . From the points of the center line C obtained by Algorithm 1 (in Appendix), parameterization with respect to s is obtained. Here s is an equi-distant discrete set of values defined along the center line. These points are later used to analyze a discrete set of longitudinal planes for parameterization of an AAA. A smooth approximation function is generated based on a basis function $\phi^i(s)$ [28] multiplied with the set of points C as in Eq. (1.1).

$$\rho(s) = \sum_{i=1}^m \phi^i(s) C(i), \quad (1.1)$$

where m is the total number of discrete points of the center line generated by Algorithm 1 of Appendix .

By applying Algorithm 1 to the four point cloud data sets $D_{scan} = \{D_1, \dots, D_4\}$, we obtained observations of center lines $\{\bar{\rho}(s, t_i) | s \in \mathcal{S}_i\}$, where $\forall i \in \mathcal{I} := \{1, \dots, 4\}$.

1.2.3 Observations of AAA Surfaces from the Data

The surface data is then parameterized with respect to the center line by defining a function $r : \mathcal{S} \times \Theta \rightarrow \mathbb{R}_{>0}$, where $\mathcal{S} := [0, z_{max}]$ and $\Theta := [0, 2\pi]$ with the input coordinate system of $s \in \mathcal{S}$ as the travel length in mm along the center line and $\theta \in \Theta$ as the angle in radians. The output of r is given as the distance from the center line to the point on the surface at a given set of input coordinates (s, θ) . A visualization of this coordinate system is shown in Fig. 1.1. Therefore, in this chapter, an AAA is modeled by $r(s, \theta)$ with respect to $\rho(s)$, where $s \in \mathcal{S}$ and $\theta \in \Theta$. The detail information regarding how to obtain a noisy version of

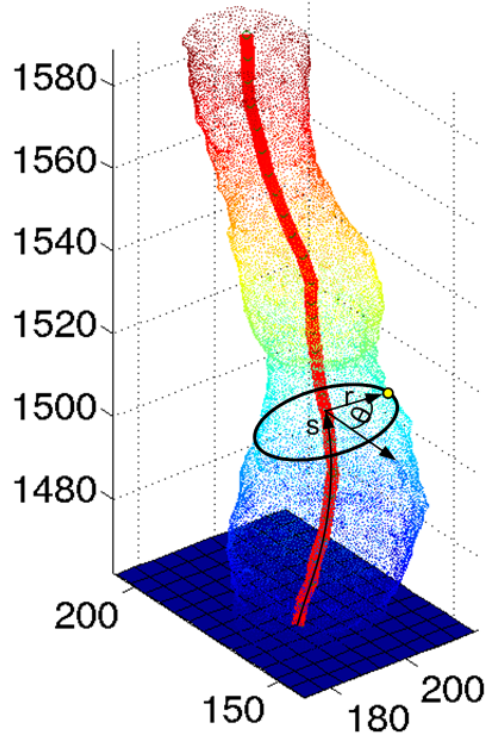


Figure 1.1: Parametrized axis system $r(s, \theta)$ (black) and center line (red), where s is the travel length in mm along the center line and $\theta \in \Theta$ is the angle in radians. The output of r is given as the distance from the center line to the point on the surface. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this thesis.

$r(s, \theta)$ from the point cloud data D is given in Appendix and summarized in Algorithm 2 of Appendix . The outputs of Algorithms 1 and 2 from the data set D are denoted as $\bar{\rho}(s, t)$ and $\bar{r}(s, \theta, t)$, where $t \in \{t_1, \dots, t_4\}$, respectively. They are considered to be observations obtained from a small number of sampling times, e.g., the limited number of CT scan images of a patient.

In summary, by applying Algorithm 2 to the four point cloud data sets $D_1, \dots, D_4$, we have obtained observations of AAA surfaces $\{\bar{r}(s, \theta, t_i) | s \in \mathcal{S}_i, \theta \in \Theta_i\}$, where $\forall i \in \mathcal{I} := \{1, \dots, 4\}$.

1.2.4 Gaussian Process Regression

In Section 1.3, we develop a spatio-temporal model of an AAA for a given observation set that is generated by the previous subsections. To this end, Gaussian process regression plays a key role in constructing a spatio-temporal model of an AAA. In this subsection, we briefly review Gaussian process regression. A Gaussian process is formally defined as follows [34].

Definition 1: A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.

A Gaussian process is completely specified by its mean and covariance functions. Let $x \in \mathcal{Q} := \mathcal{R} \times \mathcal{T} \subset \mathbb{R}^d$ denote the index vector, where $x := \begin{bmatrix} r^T & t \end{bmatrix}^T$ contains the sampling location $r \in \mathcal{R} \subset \mathbb{R}^{d-1}$ and the sampling time $t \in \mathcal{T} \subset \mathbb{R}_{\geq 0}$.

For an illustrative purpose, we consider a Gaussian process

$$z(x) \sim \mathcal{GP}(\mu(x), \mathcal{K}(x, x')).$$

In general, the mean and the covariance functions of a Gaussian process can be estimated a

priori by maximizing the likelihood function [32].

Suppose, we have p noise corrupted observations with $D_e = \{(x^{(i)}, \bar{z}^{(i)}) | i = 1, \dots, p\}$.

Assume that

$$\bar{z}^{(i)} = z^{(i)} + n^{(i)},$$

where $n^{(i)}$ is independent and identically distributed (i.i.d.) white Gaussian noise with variance σ_n^2 . $\mathbf{x}$ is defined as $\mathbf{x} = \text{col}(x^{(1)}, x^{(2)}, \dots, x^{(p)})$. The collections of the realizations $\mathbf{z} = [z^{(1)}, \dots, z^{(p)}]^T \in \mathbb{R}^p$ and the observations $\bar{\mathbf{z}} = [\bar{z}^{(1)}, \dots, \bar{z}^{(p)}]^T \in \mathbb{R}^p$ have the Gaussian distributions

$$\mathbf{z} \sim \mathcal{N}(\mu(\mathbf{x}), K(\mathbf{x})), \quad \bar{\mathbf{z}} \sim \mathcal{N}(\mu(\mathbf{x}), K(\mathbf{x}) + \sigma_n^2 I_p),$$

where $K(\mathbf{x}) \in \mathbb{R}^{p \times p}$ is the covariance matrix of $\mathbf{z}$ and is obtained by $K_{ij}(\mathbf{x}) = \mathcal{K}(x^{(i)}, x^{(j)})$ and $I_p \in \mathbb{R}^{p \times p}$ is the identity matrix. We can predict the value z_* of the Gaussian process at a point x_* [34] as

$$z_* | D_e \sim \mathcal{N}(\mu_*(\mathbf{x}), \sigma_*^2(\mathbf{x})), \quad (1.2)$$

where the predictive mean $\mathbb{E}(\mathbf{z} | D_e)$ is

$$\mu_*(\mathbf{x}) = \mu(\mathbf{x}) + k^T(\mathbf{x}) \left(\mathbf{K}(\mathbf{x}) + \sigma_n^2 I_p \right)^{-1} (\bar{\mathbf{z}} - \mu(\mathbf{x})) \quad (1.3)$$

and the predictive variance is given by

$$\sigma_*^2(\mathbf{x}) = \text{Var}(z_* | D_e) = \sigma^2 - k^T(\mathbf{x}) \left(\mathbf{K}(\mathbf{x}) + \sigma_n^2 I_p \right)^{-1} k(\mathbf{x}). \quad (1.4)$$

Here $k(\mathbf{x}) \in \mathbb{R}^p$ is the covariance matrix between $\mathbf{z}$ and z_* obtained by $k_j(\mathbf{x}) = \mathcal{K}(x^{(j)}, x_*)$ and $\sigma^2 = \mathcal{K}(x_*, x_*) \in \mathbb{R}$ is the variance at x_* .

It can be seen from Eqs. (1.3) and (1.4) that the calculation of both the predictive mean and predictive variance requires the inversion of covariance matrix whose size depends on the number of observations p , i.e., its complexity is $\mathcal{O}(p^3)$. Hence a drawback of Gaussian process regression is computational complexity. A large p makes it impossible to compute Eqs. (1.3) and (1.4) using all data points. To overcome the limited computation resource, a number of approximation methods have been proposed. For instance, the sparse greedy approximation method [35], the Nystrom method [36], the informative vector machine [37], the likelihood approximation [38], and the Bayesian committee machine [39] have been employed for different problems. In particular, it has been proposed that spatio-temporal Gaussian process regression can be applied to truncated observations including only measurements near the position and time of interest [31]. To justify prediction based on only the most recent observations, a similar argument has been made in [40] in the sense that the data from the remote past do not change the predictors significantly under the exponentially decaying correlation functions. In this chapter, to cope with computation complexity, we will also use local observations near the point of interest when we compute the prediction of that target point.

1.3 Spatio-temporal Modeling of an AAA

From now on, we explain how to model the evolution of an AAA of a patient based on the limited data set of CT scan images such that estimation (or prediction) and the error variance of an AAA can be computed for any given time (including future time). In this section, we

will use noisy observations of center lines $\{\bar{\rho}(s, t_i) | s \in \mathcal{S}_i\}$ and surfaces $\{\bar{r}(s, \theta, t_i) | s \in \mathcal{S}_i, \theta \in \Theta_i\}$, where $\forall i \in \mathcal{I} := \{1, \dots, 4\}$ computed from the four point cloud data sets $D_1, \dots, D_4$ as described from Sections 1.2.2 and 1.2.3.

1.3.1 Spatio-Temporal Modeling of the Center Line

The x , y and z coordinates of center lines for previous times are independent of each other both in spatial and temporal directions. Hence, the prediction framework of the center line of an AAA utilizes three independent zero mean Gaussian processes as:

$$\rho_x(s, t) \sim \mathcal{GP} \left(0, \mathcal{K}_x(s, t, s', t'; \Phi_x) \right),$$

$$\rho_y(s, t) \sim \mathcal{GP} \left(0, \mathcal{K}_y(s, t, s', t'; \Phi_y) \right),$$

$$\rho_z(s, t) \sim \mathcal{GP} \left(0, \mathcal{K}_z(s, t, s', t'; \Phi_z) \right),$$

where $t \in \mathcal{T}$ and $\rho_x(s, t) := \rho(s, t) \cdot e_x$, $\rho_y(s, t) := \rho(s, t) \cdot e_y$, and $\rho_z(s, t) := \rho(s, t) \cdot e_z$ are the x , y and z coordinates of the center line at time t with distance s mm from the origin. Here e_x , e_y , and e_z denote the unit vectors codirectional with the x , y , and z axes, respectively. The standard exponential kernel function is used for calculating the covariance function for each of directions:

$$\mathcal{K}_\alpha(s, t, s', t'; \Phi_\alpha) = \sigma_{f\alpha}^2 \exp \left(-\frac{|t - t'|^2}{2\sigma_{t\alpha}^2} \right) \times \exp \left(-\frac{|s - s'|^2}{2\sigma_{s\alpha}^2} \right)$$

with hyperparameters $\Phi_\alpha := [\sigma_{f\alpha} \sigma_{s\alpha} \sigma_{t\alpha}]^T$ where $\forall \alpha \in \{x, y, z\}$. $\sigma_{s\alpha}$ and $\sigma_{t\alpha}$ are bandwidths for space and time. The hyperparameters are determined by maximizing the likelihood function. The obtained hyperparameters are shown in Table 1.2. Having estimated

hyperparameters in $\{\Phi_\alpha, \forall \alpha \in \{x, y, z\}\}$ from observations $\{\bar{\rho}(s, t_i) | s \in \mathcal{S}_i\}$, where $\forall i \in \mathcal{I}$, using the covariance function form in Eq. (1.5), we can now predict the center line of the AAA for any time using Gaussian process regression illustrated in Section 1.2.4. The prediction will be given at any space and time (s_*, t_*) by the conditional expectation:

$$\hat{\rho}(s_*, t_*) := \mathbb{E}(\rho(s, t) | \{\bar{\rho}(s, t_i) | s \in \mathcal{S}_i\}, \forall i \in \mathcal{I}).$$

Since the point-wise variance obtained in each coordinate dimension is independent, the uncertainty envelop obtained across each point of center line is an ellipsoidal. The predicted center line for the time of fourth scan t_4 using center line data of first three scans along with the original center line obtained is shown in Fig. 1.2.

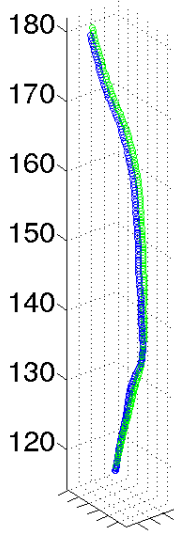


Figure 1.2: The predicted center line for fourth scan using first three scans. The predicted center line is shown in green and the original center line is in blue.

1.3.2 AAA Surface Prediction

In this section, the AAA surface $r(s, \theta, t)$ is modeled by using Gaussian process regression using observations of AAA surfaces $\{\bar{r}(s, \theta, t_i) | s \in \mathcal{S}_i, \theta \in \Theta_i\}$, where $\forall i \in \{1, \dots, 4\}$.

We assume that the AAA surface parameter r is a Gaussian process, i.e.,

$$r(s, \theta, t) \sim \mathcal{GP}(\mu_r, \mathcal{K}(s, \theta, t, s', \theta', t'; \Psi)),$$

where the covariance function $\mathcal{K}(s, \theta, t, s', \theta', t'; \Psi)$ with a hyperparameter vector $\Psi := [\sigma_f, \sigma_s, \sigma_\theta, \sigma_t]^T$ is calculated using the kernel function [41] given as:

$$\mathcal{K}(s, \theta, t, s', \theta', t'; \Psi) = \sigma_f^2 \exp\left(-\frac{|s - s'|^2}{2\sigma_s^2}\right) \exp\left(-\frac{1 - \cos(\theta - \theta')}{2\sigma_\theta^2}\right) \exp\left(-\frac{|t - t'|^2}{2\sigma_t^2}\right), \quad (1.5)$$

where σ_f^2 presents the range on which r varies vertically for a fixed point. Again, σ_s and σ_t are bandwidths in space and time. The effect of a bandwidth can be illustrated as follows. If r dose not change much in time at a spatial point, σ_t would be large and strong correlation would be reflected in the corresponding entries of the covariance matrix. The torus function for θ in Eq. (1.5) ensures that the covariance factor contributed by θ takes the highest value when $\theta - \theta' = 2N\pi$, where $N \in \mathbb{Z}$ and the lowest value when $\theta - \theta' = (2N + 1)\pi$, where $N \in \mathbb{Z}$.

The hyperparameters in Ψ are calculated by maximizing the likelihood function [32]. The estimated parameters are given in Table 1.2. Using the covariance function in Eq. (1.5) with the estimated hyperparameters plugged-in, the Gaussian process regression can be performed as discussed in Section 1.2.4. In particular, the prediction r can be made at any input point

and time (s_*, θ_*, t_*) given by the conditional expectation

$$\hat{r}(s_*, \theta_*, t_*) := \mathbb{E}(r(s_*, \theta_*, t_*) | \{\bar{r}(s, \theta, t_i) | s \in \mathcal{S}_i, \theta \in \Theta_i\}, \forall i \in \mathcal{I}).$$

Table 1.2: Hyperparameters For Surface Prediction

Gaussian Process	Hyperparameters	Estimated Values
Surface Prediction	$\Psi := [\sigma_f, \sigma_s, \sigma_\theta, \sigma_t]$	[23.21, 62.3, 1.72, 336]
Center line Prediction	$\Phi_x := [\sigma_{fx}, \sigma_{sx}, \sigma_{tx}]$	[12.3, 63.6, 336]
	$\Phi_y := [\sigma_{fy}, \sigma_{sy}, \sigma_{ty}]$	[28.4, 67.4, 336]
	$\Phi_z := [\sigma_{fz}, \sigma_{sz}, \sigma_{tz}]$	[123.5, 137.5, 336]

As discussed in Section 1.2.4, prediction at the target spatio-temporal point is made using the points nearest to its spatial locations in previous scans. For our case, each scan consists of 66600 distinct spatial points with $s \in \mathcal{S} := \{0, 1, \dots, 184\}$ and $\theta \in \Theta := \{0, 1, \dots, 360\}$. The subset of data used for making prediction feasible at a single point $(\bar{s}, \bar{\theta}, t_i)$ for case j (to be defined shortly), where $j \in \mathcal{J}$, is given by

$$\{\{\bar{r}(\bar{s}, \bar{\theta}, t_i) | \bar{s} \in N_i(\bar{s}) \subset \mathcal{S}_i, \bar{\theta} \in \Theta_i\}, \forall i \in \mathcal{I}_j\}, \quad (1.6)$$

where $N_i(\bar{s})$ is the set of local indices near $\bar{s}$ at time index i . The cardinality of this index set will determine the computational complexity of Gaussian process regression. For example, if its cardinality $|N_i(\bar{s})|$ is less than N_{max} , then the complexity will be less than $\mathcal{O}(N_{max}^3)$ as can be obtained from Eqs. (1.3) and (1.4). As discussed previously, similar approaches using local observations to reduce the complexity were proposed in [31] and [40]. Xu *et al.*, [31] showed that the quality of prediction based on truncated observations does not deteriorate

much as compared with that of prediction based on all data points. For further details, the reader is referred to Theorem 3.4 on the error analysis in [31].

1.4 Case Study and Results

In this section, using the observations $D_{scan} = \{D_1, \dots, D_4\}$ of the four scan images, we formulate problems for three cases in order to illustrate the effectiveness of our approach in realistic scenarios. To validate our approach, the main idea is that we pretend that one data set D_i not available and make its prediction only using remaining data sets $\{D_j : j \neq i\}$. We then compare the prediction results with the existing true data set D_i . We summarize our three cases as follows.

- *Case 1: (Extrapolation)* For the first case, we use D_1 and D_2 taken at t_1 and t_2 to predict the AAA at t_3 . The third scan is D_3 which is available but we pretend it is not. In particular, prediction results are obtained for the center line $\{\hat{\rho}(s, t_3)\}$ and the AAA surface $\{\hat{r}(s, \theta, t_3)\}$. The reconstructed AAA is then compared with the true third scan image of D_3 . The results of case 1 are shown in Fig. 1.3 and Fig.x 1.4.
- *Case 2: (Extrapolation)* For case 2, the first three scans (D_1, D_2, D_3) are used to get prediction results at time of the fourth scan image, i.e., t_4 . Similar to the previous case, both center line and surface predictions are made and the three dimensional AAA surface is reconstructed to form the vessel. The results of this case are shown in Fig. 1.5 and Fig. 1.6.
- *Case 3: (Interpolation)* For the third case, interpolation is performed using the first two D_1, D_2 and fourth D_4 scans taken at t_1, t_2 and t_4 respectively. The interpolation results are obtained at the time of third scan image, i.e., at t_3 . Like previous

cases, reconstruction is performed using the interpolated center line $\{\hat{\rho}(s, t_3)\}$ and the interpolated parameterized surface $\{\hat{r}(s, \theta, t_3)\}$. The results of this case are shown in Fig. 1.7 and Fig. 1.8.

Considering Eq. (1.6), the new data sets for cases 1, 2, 3 can be systematically organized by $\mathcal{J} = \{1, 2, 3\}$, $\mathcal{I}_1 = \{1, 2\}$, $\mathcal{I}_2 = \{1, 2, 3\}$, and $\mathcal{I}_3 = \{1, 2, 4\}$.

The hyperparameter vectors Ψ and $\{\Phi_\alpha\}$, where $\alpha \in \{x, y, z\}$, estimated by maximizing the likelihood function for the given data, are shown in Table 1.2.

For each case, the root mean square error (RMSE) is calculated by comparing the two functions of the original and predicted parameterized AAA surfaces. For a fixed time, the original surface is represented as $\bar{r}(s, \theta)$ whereas the predicted surface is represented as $\hat{r}(s, \theta)$. The two surfaces are grid-aligned and the RMSE is calculated for all values of s and θ as follows.

$$RMSE = \sqrt{\frac{1}{n \times k} \sum_{i=1}^n \sum_{j=1}^k (\bar{r}(s_i, \theta_j) - \hat{r}(s_i, \theta_j))^2} \quad (1.7)$$

The RMSE and maximum error between collections of grid points on surfaces for all cases are given in Table 1.3.

Table 1.3: Error measures in Prediction using data of Patient B

Case Number	RMSE (mm)	Maximum Error
Case 1	3.2	10.5
Case 2	2.05	6.9
Case 3	1.6	4.9

For case 1, Fig. 1.3 shows the prediction of both parameterized surface and reconstructed surface of an AAA using the predicted center line at time t_3 . As can be shown, the prediction results quite accurately match with the original surface. With increase of longitudinal data

as in case 2, the prediction results further improve as can be shown in Table 1.3. The prediction results along with visualization of original data for time t_4 for case 2 are shown in Fig. 1.5. In case 3, which used data of three CT scans for interpolation, the results obtained were the best of three cases. This is expected in a sense that nonparametric regression (such as Gaussian process regression) performs better in interpolation than in extrapolation (prediction at future time). The interpolated surface of an AAA at time t_3 along with data from original scan for case 3 is shown in Fig. 1.7. Table 1.3 summarizes these results using the RMSE. It shows a decrease in error from 3.2 mm to 2.05 mm with increase of one scan and the best results for interpolation with RMSE going down to 1.6 mm.

1.5 Discussion

In this section, we discuss possible utility and limitations of our approach along with future research directions.

1.5.1 Decision Making via Prediction and its Confidence Region

The major possible utility of our algorithms is to help clinicians in conducting medical treatment of an AAA (such as monitoring, open surgery or endovascular repair) by providing the predicted AAA (at future time) and its confidence region generated from the limited number of available CT scan images. The results from our case study showed excellent performance of our algorithms under three different cases.

Prediction error variances for predicted values can be computed using Eq. (1.4), which is one of the main advantages of using Gaussian process regression to model AAAs. Using Eq. (1.4), we can compute the confidence regions. For a clear visualization, let us present

a confidence region for the predicted parameterized AAA surface. The surface predicted at t_4 for case 2 along with point-wise upper and lower 90%-confidence intervals are shown in Fig. 1.9. The confidence regions can be straightforwardly computed for three dimensionally reconstructed AAAs. In this way, uncertainty quantification in predicted AAAs, however, is readily available by correctly taking into account all uncertainties, for example, available CT images, different observation noise (or resolution) levels in CT images, and patient-specific estimated hyperparameters in an empirical Bayes method. This capability of gauging uncertainty in the predicted AAAs, however, is not available in standard G&R computational models [22–25]. Again, confidence regions on predicted AAAs will be very useful in making clinical decision in order to gauge the level of confidence in any decision made.

1.5.2 Scheduling of CT Scans

The number of scans and the time difference between scans are influential in generating a good quality prediction of an AAA at a particular time. Since a large time difference, in the prediction phase, would result in little correlation, higher uncertainties in final prediction would occur. In general, a large number of scans for a patient is also desirable for better quality of both hyperparameter estimation and AAA prediction. This implies that the confidence region is a function of the scanning times and other parameters such as resolution, noise levels etc. Therefore, given all other parameters and previous CT scans of a particular subject, the next CT scan can be scheduled in order to meet a desired level of prediction quality by calculating its confidence region. Note that once hyperparameters are fixed, prediction and its confidence region can be calculated at any future time as illustrated using Eqs. (1.3) and (1.4).

1.5.3 Hyperparameters as Possible Feature Vectors

The hyperparameters estimated for the given data are shown in Table 1.2. For the surface of an AAA, σ_f in Eq. (1.5) is an indication of the range on which the radius r varies for a given input point. The hyperparameter σ_s in Eq. (1.5) is the scaling factor in direction of center line s and captures the correlation structure of the surface along s . For example, a high value for σ_s implies that the AAA surface varies smoothly whereas a lower value indicates that the surface has high variance in direction of s . Similarly σ_θ in Eq. (1.5) is the scaling factor for θ and σ_t in Eq. (1.5) is the temporal scaling factor in the covariance structure.

The hyperparameter vector can be viewed as a feature vector that may encode information of the AAA evolution. The hyperparameters estimated for the regression provide a unique patient-specific feature vector which captures both the temporal and spatial variation patterns across and around the length of AAA surface. Collective feature vectors obtained from more patients could be useful in building a classification module capable of detecting patients with imminent danger of rupture [42]. In the presence of more longitudinal data, an estimation of the temporal hyperparameter would also be a guide for specifying the ideal difference at which CT scans of AAA should be conducted for a specific patient.

1.5.4 Limitations and Future Research Directions

Our current method presented in this chapter is based on an empirical Bayes method where estimators for uncertain values such as hyperparameters and center lines are plugged in (as approximation) instead of integrating out the uncertainties in such variables (as in a fully Bayesian way). Hence, uncertainties in such variables are not fully accounted while gauging confidence regions. However, prediction error variances in center lines are small and can be

easily accounted in confidence regions of predicted AAAs. Gaussian process regression is robust to selection of hyperparameters. It is a common practice that hyperparameters are obtained *a-priori* by maximizing the likelihood function as an empirical Bayes fashion [34]. We have justified our approach of using an empirical Bayes method by showing excellent prediction results with respect to true AAAs that were not used in training data for our case study in Section 1.4. The fully Bayesian approach using Gaussian process regression with an uncertain covariance function is computationally expensive. This will add much more complexity to the current one of $\mathcal{O}(n)$ with n observations. In addition, prior distributions on uncertain variables need to be carefully selected. For further information, the reader is referred to [43, 44]. Hence, a future research direction is to develop a fully Bayesian version of our proposed scheme taking into account uncertainties in hyperparameters and center line prediction. Given the excellent results from our current method using an empirical Bayes method, even if a fully Bayesian approach is used, we won't expect significantly better performance resulted. Nonetheless, it can provide a complete solution to our proposed formulation without any approximation used as in empirical Bayes methods.

As can be seen from the results in Table 1.3, the interpolation results (case 3) are better than those of extrapolation (cases 1 and 2). It could be expected that the quality of the predicted AAA will decrease as the prediction time horizon increases. This is more eminent in our current formulation due to the fact that the nonparametric regression technique is used without inclusion of the G&R computational model [22–25]. Our approach with inclusion of the G&R computational model will be a computationally and theoretically challenging task given the computational complexity of the model and its unknown input parameters. However, a well-adopted computation model structure will provide a constraint in space and time, which will help in reducing the size of the confidence region of the predicted AAA

at future time. Therefore, the incorporation of the computational model in our Bayesian framework shall be our future research direction.

1.6 Conclusion

In this chapter, we formulated the AAA modeling and its growth using patient-specific CT scan image data in a purely statistical framework. As part of the work, a unique visualization of an aneurysm is provided using a surface parameterization in $r(s, \theta, t)$ coordinate system with respect to a center line of $\rho(s, t)$ at time t . Using the proposed methodology and available CT scan data, the prediction of an AAA can be made for any time using truncated Gaussian process regression. The results of the case study showed excellent performance of our algorithms when they are compared to the true CT scan images. To the best of our knowledge, this is the first study that predicts the AAA growth using available (patient-specific) CT scan data in a statistical perspective allowing uncertainty quantification in the predicted AAA. In doing so, it provides some interesting insights along with limitations of such models for studying the nature of AAA growth. Possible utility and limitations of our approach along with future research directions have been discussed. With advances in computing technology and new sampling methods, the use of the Bayesian approach will have a great potential to revolutionize application of computational modeling in the treatment of vascular diseases.

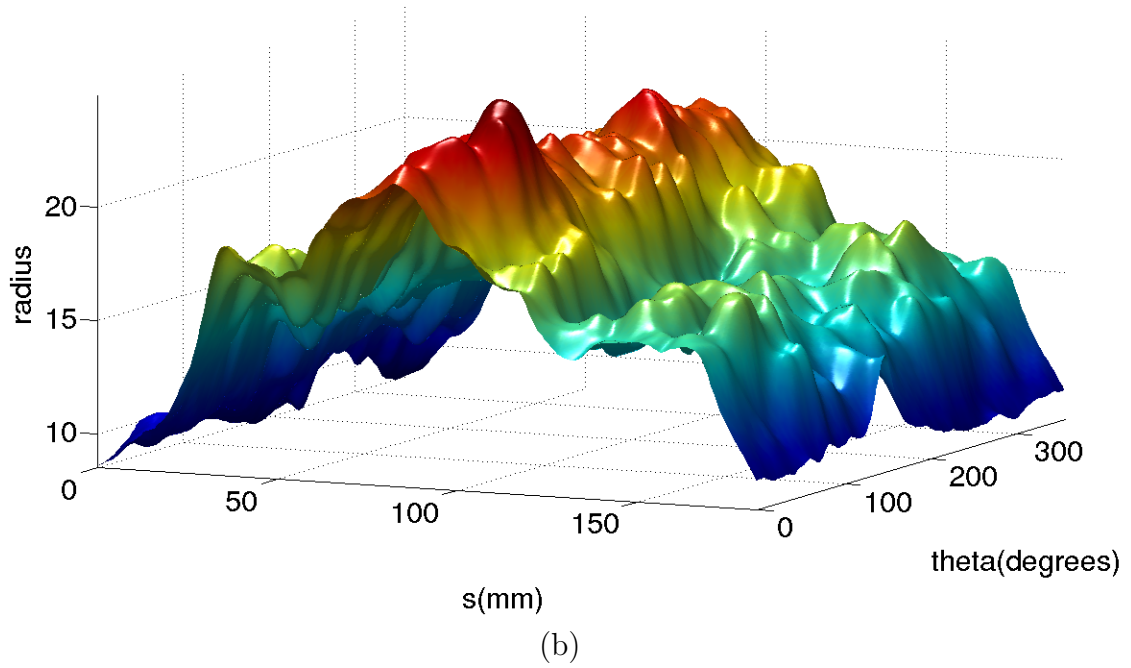
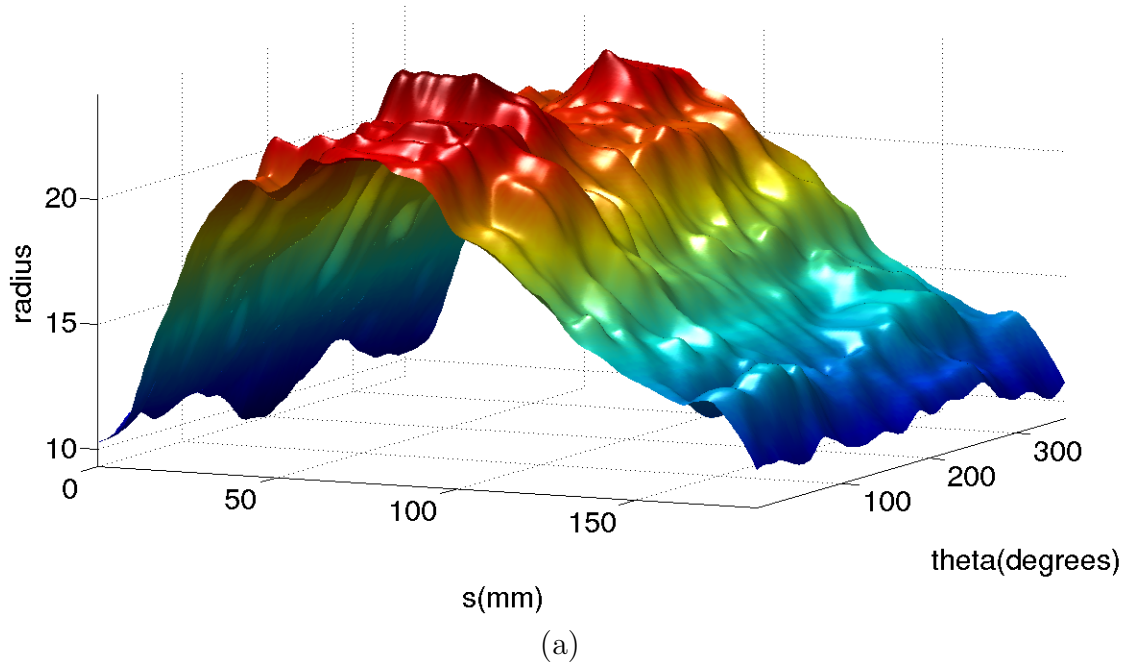
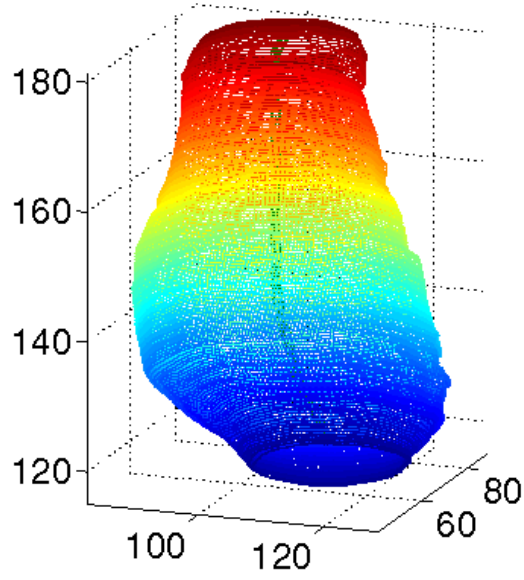
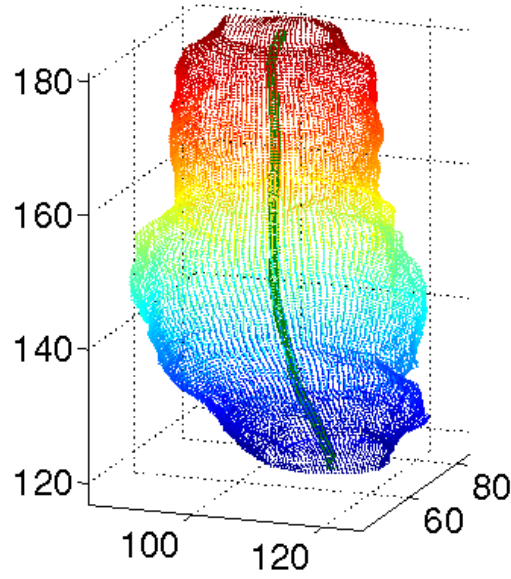


Figure 1.3: Case 1: (a) Parameterized surface using original data for time t_3 . (b) Parameterized surface using results of prediction for time t_3 .

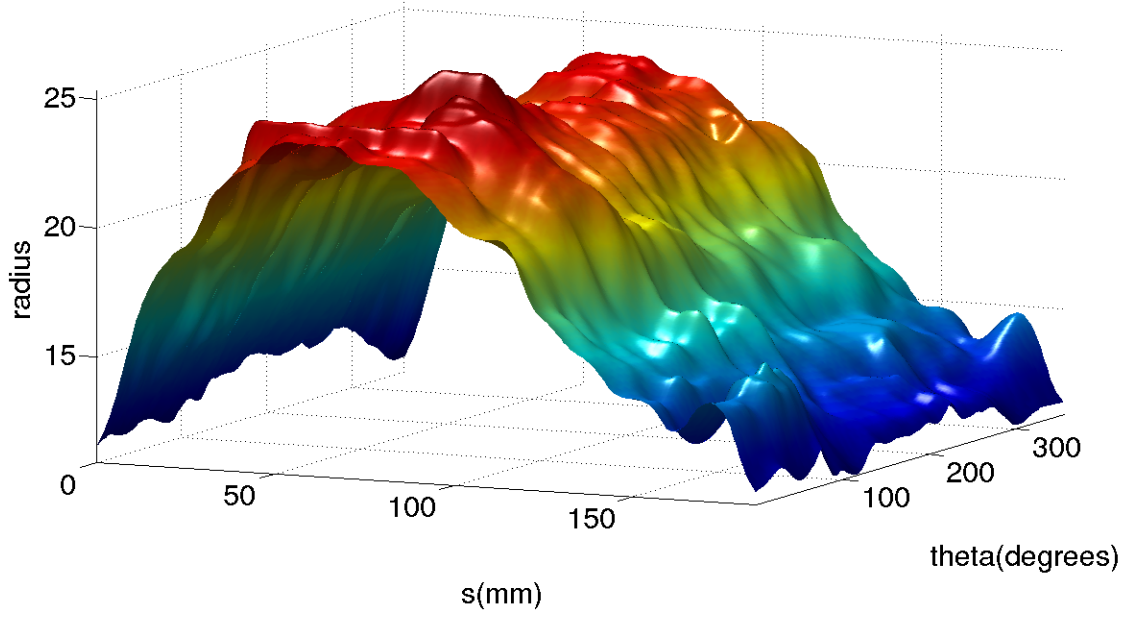


(a)

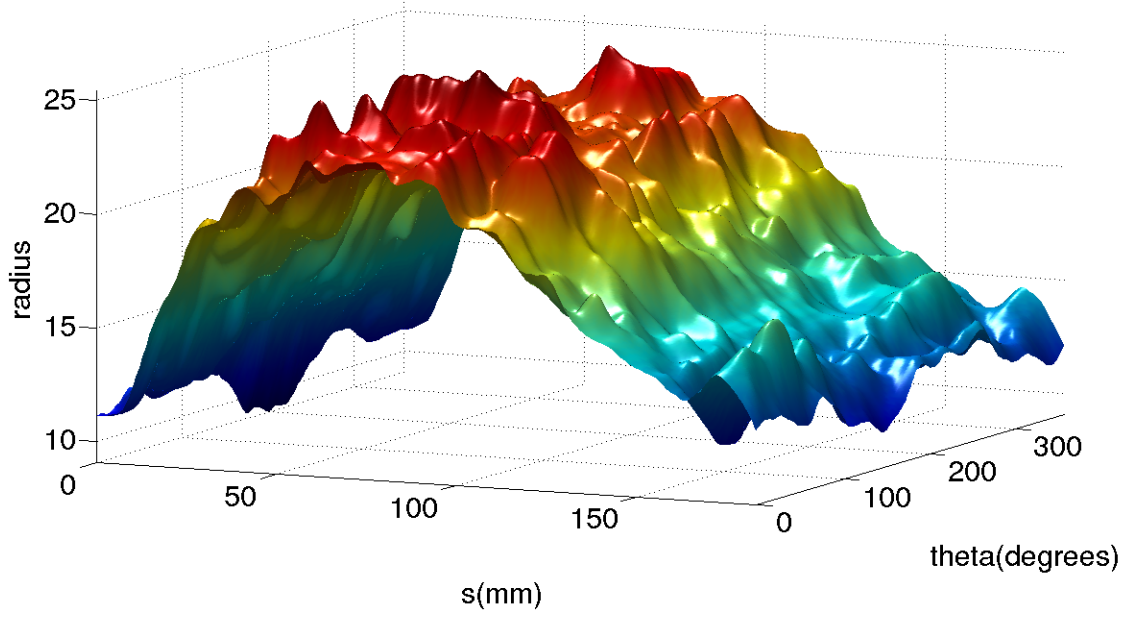


(b)

Figure 1.4: Case 1: (a) Original surface of aorta at time t_3 . (b) Reconstructed image of aorta using predicted surface and center line for time t_3 .



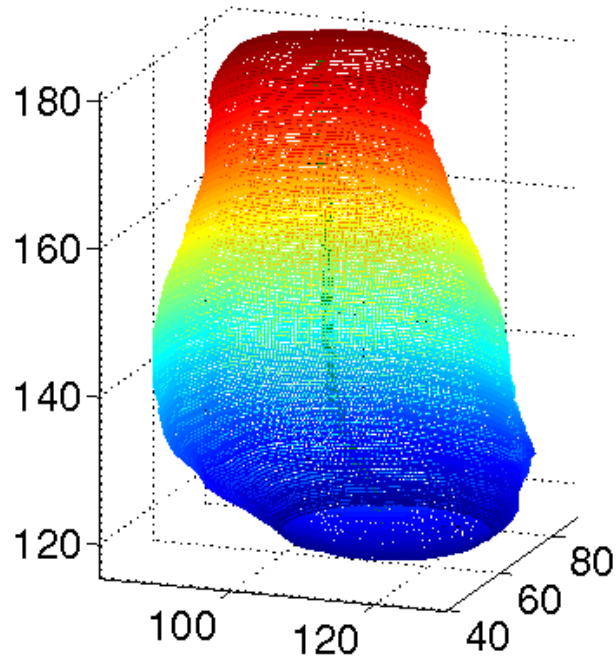
(a)



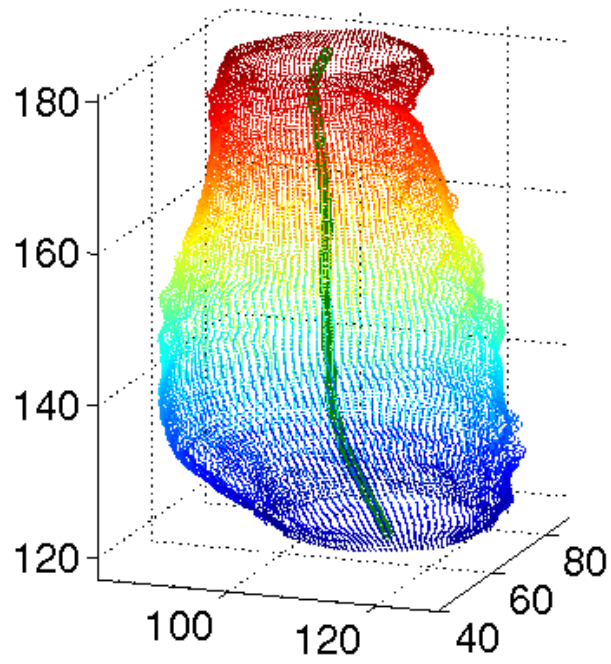
(b)

XS

Figure 1.5: Case 2: (a) Parameterized surface using original data for time t_4 . (b) Parameterized surface using results of prediction for time t_4 .



(a)



(b)

Figure 1.6: Case 2: (a) Original surface of aorta at time t_4 . (b) Reconstructed image of aorta using predicted surface and center line for time t_4 .

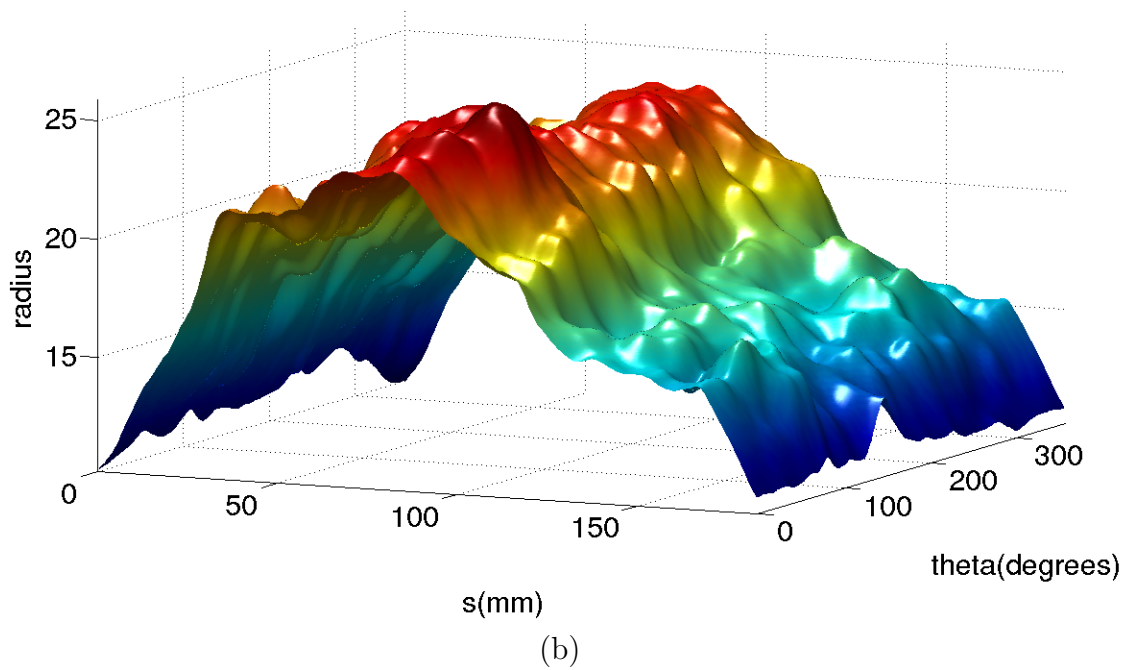
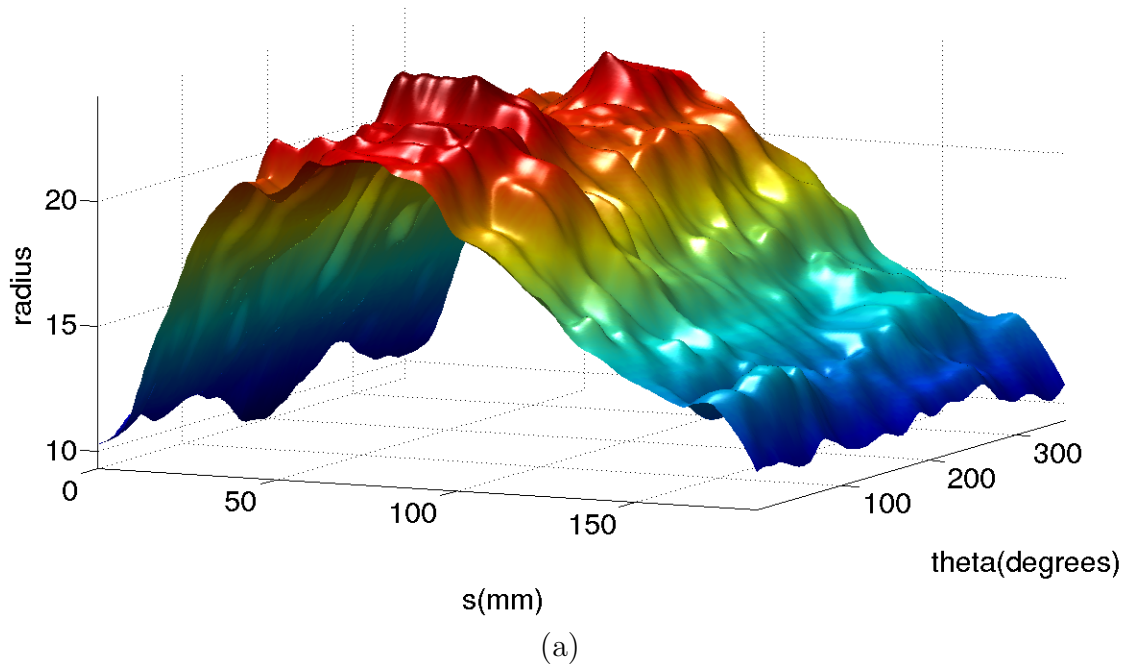
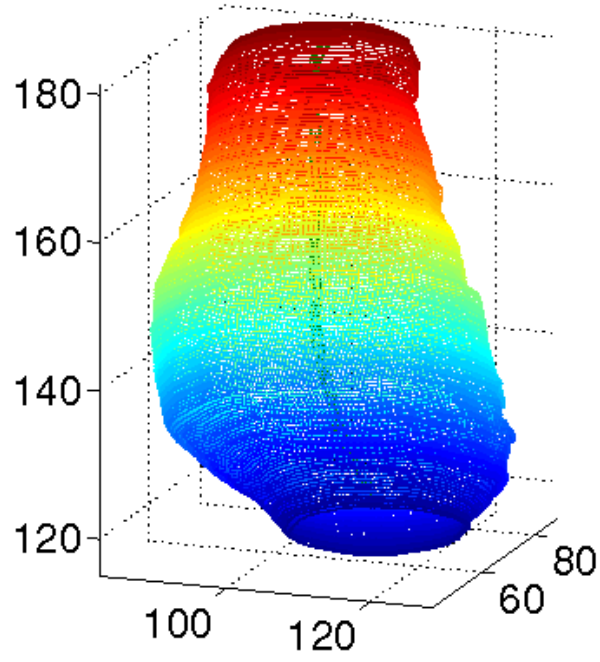
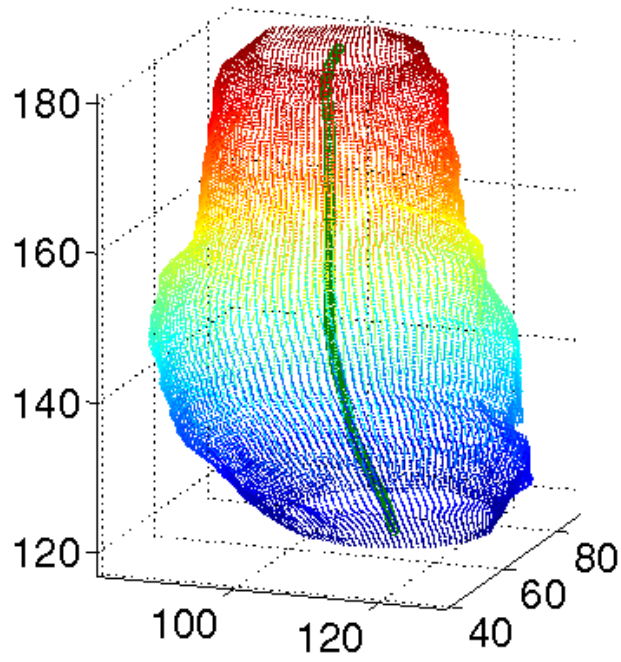


Figure 1.7: Case 3: (a) Parameterized surface using original data for time t_3 . (b) Parameterized surface using results of interpolation for time t_3 .



(a)



(b)

Figure 1.8: Case 3: (a) Original surface of aorta at time t_3 . (b) Reconstructed image of aorta using interpolated surface and center line for time t_3 .

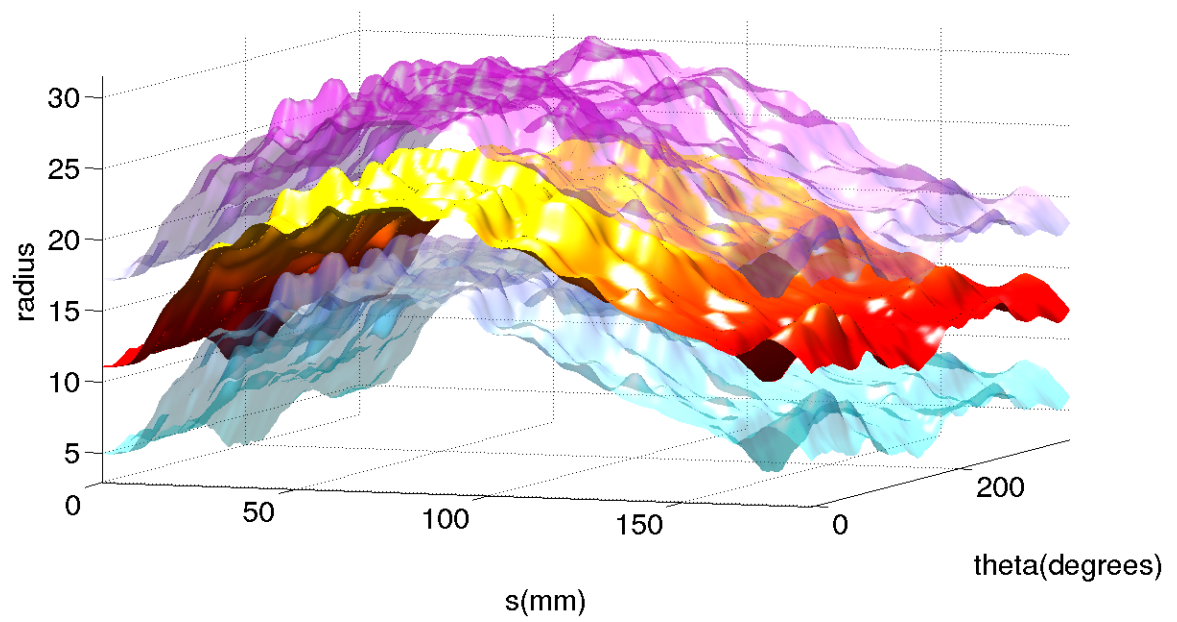


Figure 1.9: Case 2: Predicted surface (middle) with confidence intervals (up and down) at time t_4 .

Chapter 2

Temporal Modeling and Forecasting of Sentiments in Online Social Networks

Recent proliferation of online social networks (OSN) as a media for expressing opinions and sentiments has met with increasing interest from a gamut of fields. These sentiments often expressed as textual nuggets proves to be an invaluable resource for a constant assessment of policies, product review, reactionary responses, and as a feedback for improvement on strategies.

microblogging and the blogosphere have brought the sentiments expressed Sentiments on online networks are important for a variety of purposes. For instance, it is important for marketers, economists, political scientists, medical. This wide range of interest is further helped by the vast amount of data available for analysis given the microblogging and news consumption culture online.

A multi-faceted role of microblogging framework has emerged. Users provide recommendations and express sentiments while assimilating information through these diverse sources. The power and role of these networking sites has been highlighted by many studies. It has been shown that the OSNS can be used to detect World events [45], provide inference

about public health [46], predict election results [47] and play a pivotal role during emergency situations [48]. The ease of spreading information has further morphed these media sources as a platform for campaigning and modern day activism. The opinions expressed in microblogging is also shown to be highly correlated with traditional polls [49]. Thus, the public using these sources provide live feedback about major events, products and Government decisions. The e-commerce also makes use of the online feedback system for the recommendation systems [50,51]. The textual exchange has also been shown to be a source for opinion formation [52]. Therefore, discovering knowledge from this information is imperative in policy making, economists and political scientists. This necessitates a proper understanding and quantification of opinions and sentiments expressed through OSNS. Moreover, sentiments expressed through web are dynamic and vary over time. Thus, capturing users temporal preference is of prime importance [53]. The quantification of expressed information is often achieved through sentiment analysis techniques [54] whereas the temporal behavior is primarily captured through aggregation over time [49]. While such models for capturing temporal behavior of sentiments are shown to be effective, they are tied with issues of scalability, sampling and quantification of sentiment classification error. The scalability issues arises due to the sheer amount of data being generated in OSNS [45]. The sampling problems are a result of *data acquisition bottlenecks* and *information analysis complexity* [55]. Furthermore, the researchers are also constrained to use certain samples based on application and quality of data. For example, it has been shown that about 40% of all tweets from Twitter feed are pointless babbles. Also, for sentiment analysis, it is necessary to identify and use samples with subjective information.

In this study, we propose a computationally efficient mathematical model for characterizing temporal behavior of sentiments. The proposed model incorporates the inherent problems

associated with traditional aggregating schemes for temporal modeling of sentiments. The model can be used to study the changes in opinions, predict agitation in online social communities and forecast success mark of a product, campaign or political candidate. The model proposed is both flexible and extensible. It can be used with any classification scheme and easily be extended to include spatial forecasting. As a case study, we use the model to analyze the sentiments during the presidential election day 2012 of United States. The textual opinions were collected from “Twittersphere” as tweets about the two main Presidential candidates- Obama and Romney.

2.1 Related Work

Studies on temporal evolution of sentiments has been carried out in numerous field. Since all sentiment analysis and classification methods in literature comes with some classification error, it is desirable that the temporal model be flexible to incorporate it. Also, for building a real time sentiment analysis and forecasting system, the issue of scalability and sampling come into play. Acquiring textual data samples is usually constrained with rate limitations through OSNS. For example, Twitter APIs have a cap for tweet retrieval through rate limitation. Moreover not all samples acquired are subjective opinions and hence cannot be used. Therefore, for building any real time sentiment forecasting model, there is no control over usable temporal samples.

2.1.1 Sentiment Analysis

Sentiment analysis is a prolific and growing research field. Given its profitable market and wide applicability, keen interest has been shown in this field. Several studies have been conducted for improving the methods of sentiment mining and classification. An exhaustive survey prior to 2008 is [54]. The process generally consists of identifying subjective sentences [56] [57], followed by feature extraction [58–60], polarity assignment [61] and classification. The Recent studies show that the application dependent variation and complexity associated with sentiment analysis provides further room for improvement. Techniques for refined feature extraction have been suggested. [62] compensated for frequency bias of discriminative features using a frequency based weight penalizing prior on the regularization process in elastic net framework. The case for application dependent variation is made through extraction of target dependent sentiment expression from Twitter data [63]. The study also incorporates and assigns polarity to the *informal* language of tweets. Another interesting problem in sentiment analysis is acquiring training data. Several studies manually annotate labels for a small sample from data [48, 52] which is quite expensive. Some other studies make use of tag words like emoticons [64, 65] for labeling of training set. Zhang et al. (2011) use a lexicon-based method for performing sentiment classification and then applied a supervised classifier for improving the recall by using the training examples provided in the previous lexicon-based approach.

2.1.2 Temporal Modeling

The opinions and sentiments expressed in Social networking sites are dynamic and vary over time. The importance of capturing user’s temporal preferences has been discussed

and recommendation for user specific time scale is made [53]. A number of studies use aggregation of text sentiments for time series modeling, behavior of stock market has been tied with emotions expressed in blogs [66,67]. Temporal happiness in songs, blogs and about Presidents has been modeled using the same aggregation scheme [68]. Strong correlations between textual sentiments expressed in microblog messages with contemporary polling data has been shown [49] by aggregating sentiments.

A time evolving, user-specific scale for aggregating diverse data and predicting users interests has been proposed [69]. It has further been shown that changing the time segments affects both the prediction performance and local effects. In a study for event detection using Twitter, the issue of *scalability* for detection algorithms has been raised [45]. The *quality* of usable data in microblogs has been questioned. Furthermore, effects of temporal smoothing through moving average are discussed and a suggestion for an improved *stratified sampling* technique has been made [49]. The issues of *data acquisition bottlenecks* and *information analysis complexity* are raised in [55]. The data acquisition bottleneck occurs due to rate limitations on publicly available APIs (Application Programming Interfaces) whereas huge volumes of data causes complexities in information analysis. Another study shows optimal scheduling of tweets for maximizing message diffusion [70]. The case of repeating location specific diurnal patterns has been made in [71] with another study showing consistency of culture specific diurnal mood patterns.

2.1.3 Notation

Standard notation will be used throughout this chapter. Let $\mathbb{R}, \mathbb{R}_{\geq 0}, \mathbb{R}_{>0}$, and $\mathbb{Z}$ denote, respectively, the sets of real, non-negative real, positive real, and integer numbers. I_n denotes the identity matrix of size n . For column vectors $v_a \in \mathbb{R}^a, v_b \in \mathbb{R}^b$, and $v_c \in \mathbb{R}^c$,

$\text{col}(v_a, v_b, v_c) := [v_a \ v_b \ v_c] \in \mathbb{R}^{a+b+c}$ stacks all vectors to create one column vector, and $\|v_a\|$ denotes the Euclidean norm (or vector 2-norm) of v_a . $|A|$ denotes the determinant of a matrix $A \in \mathbb{R}^{n \times n}$. Let $\mathbb{E}(z)$ and $\text{Var}(z)$ denote, respectively, the expectation and the variance of random vector z . A random vector $z \in \mathbb{R}^q$, which is distributed by a multivariate Gaussian distribution of a mean $\mu \in \mathbb{R}^q$ and a variance $\Sigma \in \mathbb{R}^{q \times q}$, is denoted by $z \sim \mathcal{N}(\mu, \Sigma)$. Let $\text{Bern}(p)$ denote a Bernoulli distribution with mean value p and $B(n, p)$ be a Binomial distribution where n is the number of trials and p represents the success probability.

2.2 Proposed Method

Classification module is developed based on labeled training data and random samples are taken from the corpus for building the temporal model. The acquired samples are filtered through a selection criteria. Feature extraction followed by sentiment classification is performed on the selected data samples and a gaussian process based spatio-temporal model is formed for prediction and forecasting of sentiments.

2.2.1 Sentiment Classification

Our study mainly concerns improvement of temporal model by incorporating the classification error from sentiment analysis. Any classification scheme can be used with some parameters explained later for the model. For illustrative purposes, we use the Naive Bayes classifier with combination of unigrams and bigrams as features as the classification module. Naive Bayes is chosen since it has been shown to work well for sentiment classification based on textual features [72]. The data set is represented as $D := \{(o_1, t_1, c_1), \dots, (o_g, t_g, c_g)\}$ where o_i is the i^{th} opinion (usually expressed as a textual nugget), t_i is the time on which

that opinion was expressed and c_i is its labeled class and g is the total number of opinions expressed in the data set. Using the Naive Bayes model for classification, class c^* is assigned to an opinion o by:

$$\begin{aligned}
c^* &= \arg \max_s P(c|w_1, w_2, w_3, \dots, w_h) \\
&= \arg \max_c \prod_{j=1}^h P(w_j|s) \times P(c) \\
&= \arg \max_c \prod_{j=1}^h P(w_j|c)
\end{aligned}$$

where w_j is the selected textual feature from the opinion o and h is the total number of features selected for analysis. The conditionals of this equation are estimated using maximum likelihood estimator.

2.2.2 Error Characterization

The classification module used for sentiment classification is tied with some classification error. In the temporal models that use aggregation for quantification of these sentiments, the incorporation of these classification errors is insufficient. The classified sentiments form one of the following four cases:

$$TP := c^* = 1|c = 1 \quad \text{True Positive}$$

$$FP := c^* = 1|c = 0 \quad \text{False Positive}$$

$$TN := c^* = 0|c = 0 \quad \text{True Negative}$$

$$FN := c^* = 0|c = 1 \quad \text{False Negative}$$

For our model, we use sensitivity and specificity given in Eq. (2.1) for incorporating the classification errors.

$$\begin{aligned}\text{Sensitivity} &= \frac{TP}{TP + FN}, \\ \text{Specificity} &= \frac{TN}{TN + FP},\end{aligned}\tag{2.1}$$

Sensitivity is the proportion of the actual positives identified to the total number of positive instances in the data set. Similarly, specificity is the ratio between the actual negatives identified and the total number of negative instances in the data set. Together, they take into account both type-I (due to false positives) and type-II (due to false negatives) errors. In our case we assign a value of 1 for a positive sentiment and 0 for a negative sentiment during the classification. Sensitivity therefore translates as the success probability of correctly classifying a sentiment as positive, whereas specificity is the success probability of correctly classifying a sentiment as negative. Classification of a sentiment as either positive or negative can therefore be modeled as a Bernoulli trial with success probability $p_\alpha = \text{Sensitivity}$ for the positively identified instances and $p_\beta = 1 - \text{Specificity}$ (also known as false positive rate) for the negatively identified instances. Therefore, the classified sentiments c^* can be modelled as a Bernoulli trial as:

$$c^* \sim \begin{cases} \text{Bern}(p_\beta) & \text{if } c^* \equiv 0 \\ \text{Bern}(p_\alpha) & \text{if } c^* \equiv 1 \end{cases}\tag{2.2}$$

Each sentiment, either classified Section as positive or negative can be represented by an independent and identically distributed (i.i.d.) Bernoulli distribution as shown in Eq. (2.2).

A time unit Δt is selected as the interval on which the sentiments are aggregated so that the time line of data used is broken down into l equally spaced time intervals. The data set thus formed can be represented as $D_i = \{c_{ni}^*, \dots, c_{ni}^*\}$ where $\forall i \in \mathcal{I} := \{1, \dots, l\}$. Note here, that n which is the number of sentiments expressed in the time interval Δt might be different for each $i \in \mathcal{I}$. The summation of these sentiments leads to two separate binomial distributions for each time interval, one for the positive sentiments expressed in that interval, and the other for the negative sentiments represented as $s_{\beta i}$ and $s_{\alpha i}$ where $\forall i \in \mathcal{I} := \{1, \dots, l\}$.

$$\begin{aligned} s_{\beta i} &= \sum_{j=1}^{\beta_i} c_{ji}^* & \forall c_{ji}^* &\equiv 0, \\ s_{\alpha i} &= \sum_{j=1}^{\alpha_i} c_{ji}^* & \forall c_{ji}^* &\equiv 1, \end{aligned} \tag{2.3}$$

where α_i is the total number of positive sentiments in the i^{th} time interval and β_i is the total number of negative sentiments expressed in the interval. This leads to l binomial distributions for the data given as:

$$s_{\gamma i} \sim \text{B}(\gamma_i, p_\gamma) \quad \forall i \in \mathcal{I}, \gamma \in \{\alpha, \beta\}, \tag{2.4}$$

With an appropriate choice of Δt according to the streaming speed of data, the values of α_i and β_i become large enough to approximate the binomial distributions of Eq. (2.4) using the Gaussian distributions [73]. The accumulated sentiments $s_{\gamma i}$ and $s_{\alpha i}$ can thus be

represented as Gaussian distributions given by:

$$s_{\gamma i} \sim \mathcal{N}(\gamma_i p_\gamma, \gamma_i p_\gamma (1 - p_\gamma)) \quad \forall i \in \mathcal{I}, \gamma \in \{\alpha, \beta\}. \quad (2.5)$$

For each $i \in \mathcal{I}$, we can get an overall estimate of the sentiment by adding the two Gaussian distributions expressed in Eq. (2.5). This leads to:

$$\begin{aligned} s_i &\sim s_{\alpha i} + s_{\beta i} \\ s_i &\sim \mathcal{N}(\alpha_i p_\alpha + \beta_i p_\beta, \alpha_i p_\alpha (1 - p_\alpha) + \beta_i p_\beta (1 - p_\beta)) \end{aligned} \quad (2.6)$$

2.2.2.1 Sampling and Scalability

The issue of scalability, sampling size and strategy have been discussed [55] with comparisons of different sampling schemes and topologies. While the study suggests to incorporate both topology and user-context over *naive methods*, it is specific for information diffusion model and might not work for modeling sentiments. Moreover, in sentiment analysis, it is required to discard certain samples due to quality and subjectivity of textual information. As of recent, affects of different sampling strategies for characterizing sentiments have not been explored. In this study however, we focus on incorporating the sampling deficiency by using a predictive model for estimating sentiments at unsampled times. We suggest to select textual information from a uniform distribution with a judicious sample size and discarding any samples that fail the sentiment analysis criteria. The missing sample points are reflected in our proposed model by an increase in point wise uncertainty.

2.2.2.2 Temporal Model

The temporal model is formulated by treating the obtained data in a Gaussian process framework. The sentiments $\bar{s}(t)$ obtained in section 2.2.2 are modeled as:

$$\bar{s}(t) \sim \mathcal{GP}(0, \mathcal{K}(t, t'; \Psi))$$

The covariance function $\mathcal{K}(t, t'; \Psi)$ used to model the sentiments has the hyperparameter vector $\Psi := [\sigma_f, \sigma_t]^T$ with kernel selected as:

$$\mathcal{K}(t, t'; \Psi) = \sigma_f^2 \exp\left(-\frac{|t - t'|^2}{2\sigma_t^2}\right)$$

The hyperparameters are estimated using the method discussed in section 1.2.4. With an estimation σ_f and Equation (2.6), the error source due to classification is accounted for as modeling it as a noise element. The noise is obtained as:

$$\sigma_n^T I \sigma_n = \alpha_i p_\alpha (1 - p_\alpha) + \beta_i p_\beta (1 - p_\beta) - \sigma_f^2 I \forall i \in \mathcal{I}. \quad (2.7)$$

2.3 Eperimental setup and Data

We evaluated our model using data from Twitter using the Twitter streaming API with “Obama and Romney” as query terms. The distribution of data collected is shown in Table 2.1. As shown in Table 2.1, more than twice of the tweets were about Obama. The same pattern was observed by [74] where it is shown that twitter users discussed Obama twice as much as Romney during the time leading up to elections. The tweets were obtained from

Table 2.1: Twitter data collected

Topic	Total tweets
Obama	3740519
Romney	1680522
Combined	5421041

November 5, 2012, 6:00am to November 7, 2012, 12:00am.

2.3.1 Training Data and Feature Extraction

We acquired the training data by using the emoticons present in tweets [75, 76]. These emoticons are used as training examples because each emoticon carries a positive or negative connotation. With identification of emoticons with either positive or negative, the information can be used to obtain a labeled training data set. The emoticons used for mapping positive and negative sentiments along with the distribution of total tweets among the two candidates is shown in Table 2.2. This table shows that Obama, as being more famous on the twitsphere, obtained more opinion nuggets through tweets as compared to that of Romney. For the training set, 16000 tweets were used with equal number of positive and negative tweets. While forming the training data, the tweets were winnowed with expulsion of the tweets of following category:

- **Non-English Languages** Any tweets of language other than English were removed from the training set.
- **Dual Candidate names** Any tweets referring to both Obama and Romney were removed.
- **Dual polarity emoticons** Any tweets containing emoticons assigned with both positive and negative polarity were removed.

- **Non-subjective tweets** Tweets without adjective were removed from the training data set. This was done since we only need to analyze opinions and presence of adjectives is shown to be highly correlated with subjectivity of sentence [57].

Table 2.2: Training Data using emoticons

Sentiment	Emoticon used	Candidate	Tweets	Total tweets
Positive	:) , :} ,:D , :))	Obama Romney	25564 8877	34441
Negative	:(, :'(, :((, :@	Obama Romney	5620 2850	8470

Adjectivity and English filters were applied using WordNet [77]. For feature extraction, some of the words in tweets were also filtered:

- **Small length words:** Any word of length less than or equal to 3 was removed
- **Candidate names:** Barrack, Obama, Mitt and Romney were removed from the training set since the distribution is biased in favor of Obama and it may select one of the names as feature contributing towards favored sentiment. The bias would corrupt the final results.
- **Emoticons:** The emoticons used were also removed.
- **Retweet information:** Any information that tells about retweet by label RT was removed
- **Mentions:** Mentions of any name by hashtags and @ were also removed from the tweets of training data.
- **Website links:** Many people post website links with their tweets. The links were also removed during creation of training data.

An example of parsed positive tweet is *Not even an American but i'm hoping for whatever happens in America somehow affects Singapore too* and a negative tweet collected in this

manner is *lost the popular vote but won by the electoral vote which is sad because the 50% for just lost their voice.*

Unigram and bigram features were obtained from these tweets. After training the classifier, 40000 tweets were randomly sampled from the data set for classification. The obtained samples were checked for adjectivity and english language the same way as the training samples. The tweets were divided as Romney and Obama tweets for comparison of sentiments for each candidate. Since the rate of tweets even when using 40000 samples is 16 tweets per minute, we aggregated the sentiments into 10 minute windows. The prediction and sampling strategies are evaluated by taking sub-samples from these aggregated sentiments. After getting the sentiments, we applied Gaussian process regression with hyperparameter learning using maximum likelihood estimation. Since $\ln p(ts|\Phi)$ is generally non-convex and can have multiple maxima, the starting points for hyperparameters were selected by visual inspection of data.

2.4 Results

2.4.1 Classification Results

Classification accuracy was obtained by using 66% training and 34% test data. Table 2.3 summarizes the results obtained. The temporal analysis of data was done only on the Unigram features with Naive Bayes classifier. It was selected because of its simplicity and good results. For the sake of completeness, comparison with Support Vector Machines and use of Bigrams is also given. Table 2.1 shows that Naive Bayes classifier with unigrams as features works better of our application.

Table 2.4 shows the likelihood ratio for top features selected using Naive Bayes algorithm.

Table 2.3: Classification Results

Feature Set	Classification framework	Accuracy	Precision
Unigram	Naive Bayes	80.7%	81.2%
	SVM	77.8%	81.3%
Bigram	Naive Bayes	73.2%	74.6%
	SVM	75.4%	79.1%

It is interesting to note that most of the top features are for negative polarity.

Results of aggregated sentiments using ten minute window are shown in Figure 2.1. As can be seen, even with a ten minute window, the results are quite noisy and it is difficult to know the true temporal pattern of sentiments expressed during the election time. However, the graph shows that Obama had an overall better sentiment score as compared to Romney. The total sentiment score for Obama during the 42 hour period near elections is **6597** whereas Romney gets **4300** positive sentiments.

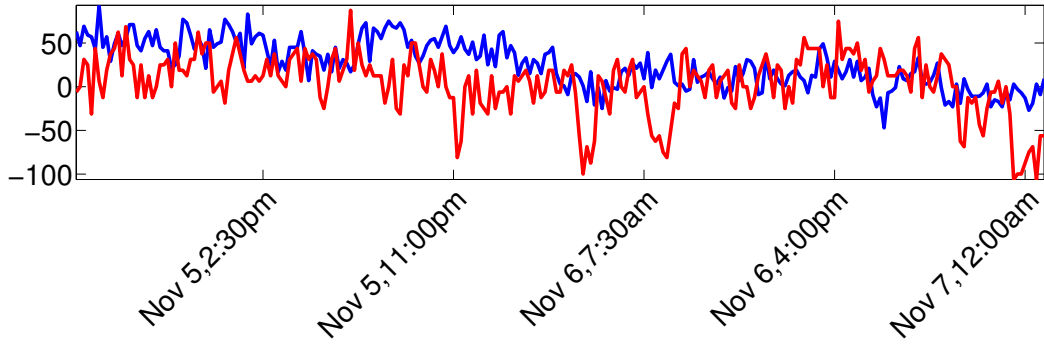


Figure 2.1: Aggregated sentiments for Obama(blue) and Romney(red)

2.4.2 Gaussian process based Temporal Model

Due to non-convex nature and multiple maximas of likelihood function, the selection of starting point for estimating hyperparameters is of prime importance. Since the prediction

Table 2.4: Most informative features

Feature Name	NEG:POS
sad	114.3:1
damn	100.8:1
hang	69.7:1
pic	48.5:1
returns	41.4:1
winning	32.3:1
chicago	29.9:1
aww	29.2:1
2013	1:28.7
lose	25.7:1
footage	24.7:1
well	1:24.6
losing	23.8:1
hilarious	1:21.8
estan	1:20.6
gone	18.4:1
cry	16.9:1
approve	1:15.6

data is low dimensional, we used visual inspection of data for selecting the initial points. The selected point for Obama's sentiments was $\sigma_f = 30$ and $\sigma_s = 10$. The local maxima was found at $\sigma_f = 23.7$ and $\sigma_s = 14.95$. For time series representing sentiments expressed about Romney, the selected initial point was $\sigma_f = 15$ and $\sigma_s = 10$ and it converged to local maxima at $\sigma_f = 12.3$ and $\sigma_s = 8.1$. Using the Gaussian process framework, the target sentiments were estimated. Figure 2.2 shows the temporal behavior of sentiments emerging from the noisy signal. This estimation is performed using sentiments obtained from only 16 hours of data instead of all 42 hours. The samples were obtained through uniform distribution and prediction using the model was performed for sentiments at unsampled time locations. The variation of point wise confidence interval in Figure 2.3 reflects the missing samples alongwith classification error noise. It is also evident from the figure that the framework models both local and global effects. The plateaus in sentiment scores are retained.

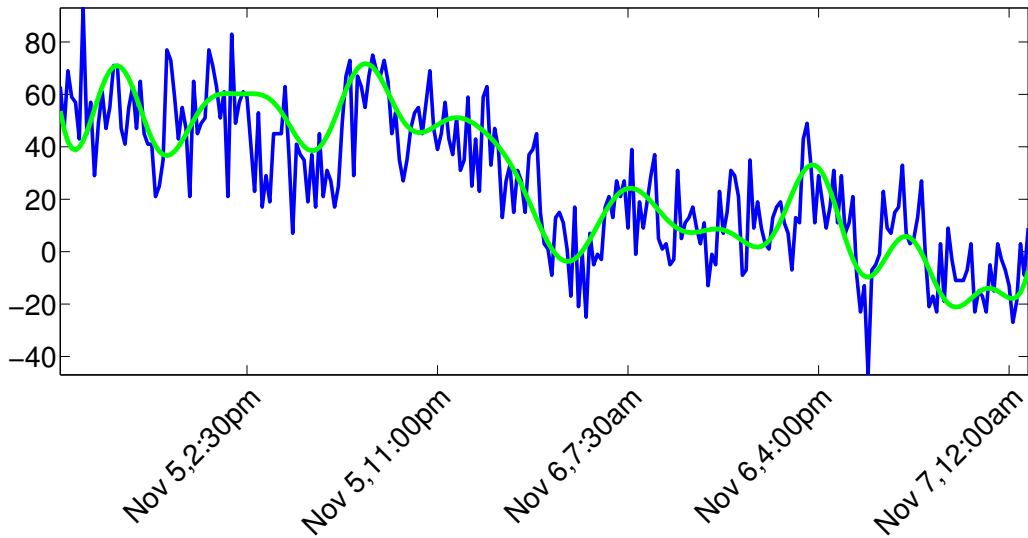


Figure 2.2: Predicted Target Sentiments(Green) with Aggregated sentiment(Blue) for Obama

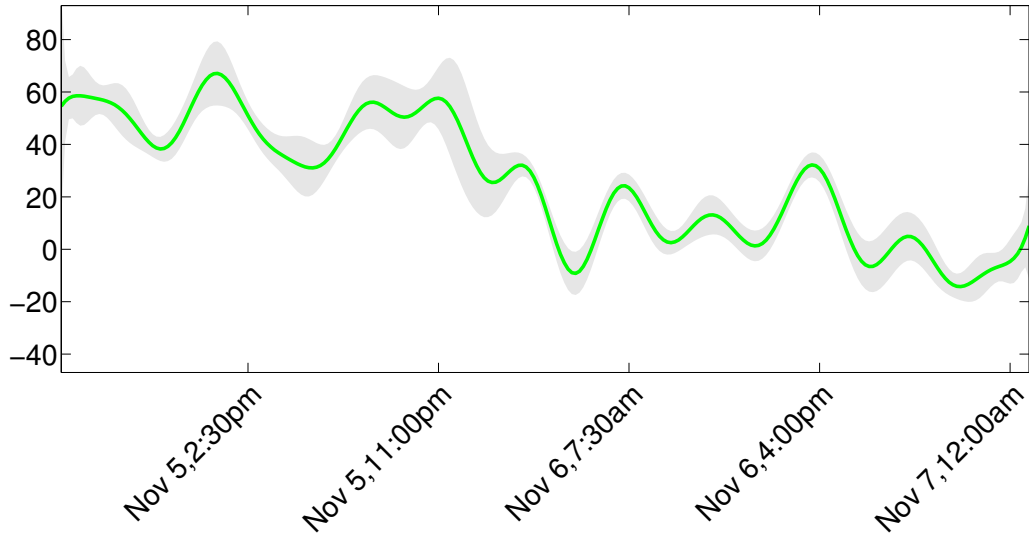
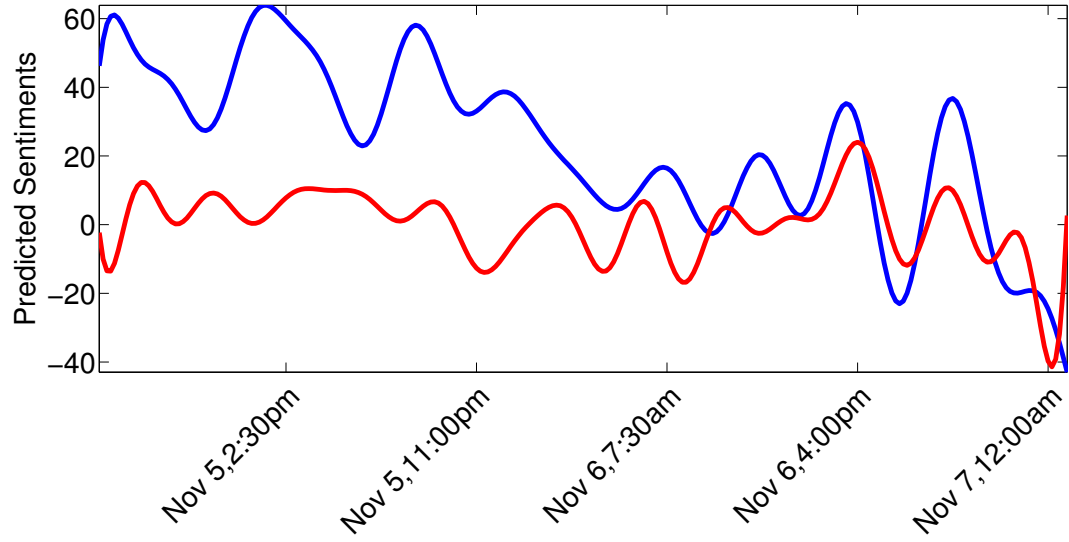


Figure 2.3: Predicted function(green) for Obama sentiments(blue) with confidence interval(grey)

The predicted target distribution of sentiments for Obama and Romney is shown in Figure 2.4. The distributions are more meaningful as compared to the noisy sentiments observed through simple aggregation.

2.4.3 Effects of Sampling

To study the effects of sample size with prediction, sentiments from labeled data of training set were obtained and the data was segmented into 255 equal time windows with each segment spanning 10 minutes. Gaussian process was used to make predictions by varying the sample size from 1 to 255. Mean square distance was calculated between predicted mean and actual data. As can be seen from Figure 2.5, the error between prediction and actual data settles down at around 60 samples. Hence, by using only 23% of data, accurate predictions can be made.



(b) Using 30 Random points of data

Figure 2.4: Predicted function(green) for Romney sentiments(red) with confidence interval(grey)

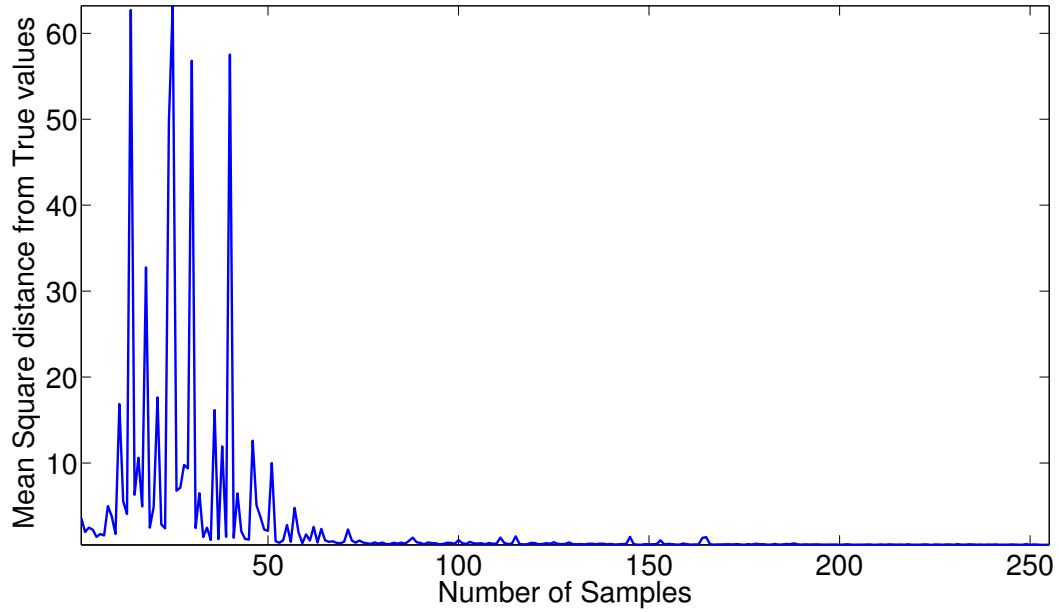


Figure 2.5: Decrease of mean square error from Original temporal sentiments with increase of Samples

2.5 Conclusion and Future Work

In this study, we have surveyed the challenges incurred during temporal modeling of sentiments. In particular, we have identified four problems; scalability, sampling, classification error and capturing both local and global phenomena. We have proposed a gaussian process framework that addresses these challenges. The extensibility of model has been discussed and mathematical formulation of spatio-temporal prediction is given. As a case study, Twitter data 42 hours prior and through the election day is used. The predicted sentiments in the temporal model have been shown to be better indicator as compared with traditional aggregating schemes. Finally, it has been shown that with only 23% of samples, high confidence for prediction can be achieved. This is still a new field of research and has many interesting problems. The error model in this study is a linear gaussian noise variable. Better stochastic modeling with deterministic parameters can be used to improve the prediction. Furthermore, Gaussian random fields can be used instead of a continuous function for decreasing the computational complexity for spatio-temporal models. This study only looked at the effects of prediction with random sampling of data. Better sampling models are desired that are optimal for sentiment modeling.

APPENDICES

Appendix A

Generation of samples of the center line

In what follows, we show how to generate a collection of finite number of samples on the center line of an AAA surface. Open ends of the vessel are required as a first step for the method. It is achieved by transversely truncating the AAA surface with truncation planes 2.5mm from the top and bottom of the vessel. The data thus obtained is a subset and is denoted as D . The center line is initialized by using the middle point of the bottom most transverse plane. A vector a is drawn between this initial guess and the point least distant from it on the AAA surface. The initial center point is then pushed in direction away from vector a by a constant δ (1mm) amount. This step is repeated for a pre-defined number of times with reduction of δ by half if the center point location hasn't changed more than 1mm in every 15 iterations. The next center line point is obtained by a linear shift from previous point in z -axis direction. An initial guess is obtained for each of the remaining center point along length of AAA by a linear shift in direction of a vector b drawn between the previous two center points. A vector c is then obtained as the projection of a onto the plane normal to b where a is calculated as explained before. The center point is translated in direction opposite to vector c by δ . This process is repeated for a pre-defined number of times with reduction of δ by half if the location of the center point hasn't changed by 1mm every fixed

iterations. The procedure is summarized by an algorithm as Algorithm 1.

Algorithm 1 Generation of Samples of Center Line

 $D = [d_\alpha(1), d_\alpha(2), \dots, d_\alpha(n)]^T \quad \forall \alpha \in \{x, y, z\}$
Output:
 $\rho(s)$
 $\triangleright$ Center line of Aorta

Algorithm:
 $l \leftarrow 1$
for all $i \in \mathcal{I}$ **do**

 if $d_z(i) = \min(d_z)$ **then**

 $b_x(l) \leftarrow d_x(i)$ when $d_z(i) = \min(d_z)$

 $b_y(l) \leftarrow d_y(i)$ when $d_z(i) = \min(d_z)$

 $l \leftarrow l + 1$

 end if
end for
 $C(1) \leftarrow \{\text{average}(b_x), \text{average}(b_y), \min(d_z)\}$
 $c_{init} \leftarrow C(1)$
for $l = 1 \rightarrow \text{MaxNumIters}$ **do**

 $d_{min} \leftarrow \min_{i \in \mathcal{I}} \|v_{D(i)/C(1)}\|$

 $a \leftarrow v_{C(1)/c_{init}}$

 $C(1) \leftarrow C(1) - \delta \times \frac{a}{\|a\|}$
 $\triangleright \delta$ is selected as 1mm–2mm

 if $l \bmod 15 = 0$ and $\|v_{c_{init}/P(1)}\| \leq 1\text{mm}$ **then**

 $\delta \leftarrow \frac{\delta}{2}$

 end if
end for
 $C(2) \leftarrow \{C_x(1), C_y(1), C_z(1) + v\}$
 $\triangleright v$ is constant

 $k \leftarrow 3$
while $C_z(k) \leq \max(d_z)$ **do**

 $b \leftarrow v_{C(k-1)/C(k-2)}$

 $C_d \leftarrow C(k-1) + v \times \frac{b}{\|b\|}$

 $c_{init} \leftarrow C_d$

 for $l = 1 \rightarrow \text{MaxNumIters}$ **do**

 $d_{min} \leftarrow \min_{i \in \mathcal{I}} \|v_{D(i)/C(k)}\|$

 $a \leftarrow v_{C_d/d_{min}}$

 $b \times \frac{a \times b}{\|b\|}$

 $c \leftarrow \frac{a \times b}{\|b\|}$

 if $l \bmod 15 = 0$ and $\|v_{c_{init}/C(k)}\| \leq 1\text{mm}$ **then**

 $\delta \leftarrow \frac{\delta}{2}$

 end if

 $C(k) \leftarrow C_d - \delta \times \frac{c}{\|c\|}$
 $\triangleright \delta$ is selected as 1mm–2mm

 end for

 $k \leftarrow k + 1$
end while
 $\rho(s) = \sum_{i=1}^m \phi^i(s) C(i)$

Appendix B

Surface parameterization

In what follows, we provide detail information regarding how to parameterize surface from the point cloud data D with respect to the calculated center line $\rho(s)$. A coordinate system N is defined to acquire longitudinal acquisition planes. The first vector defining the coordinate system $N_1(s)$ is the unit normal vector drawn between consecutive center points in $\rho(s)$. $N_2(s)$ uses the known Cartesian standard basis perpendicular to $N_1(s)$ where as $N_3(s)$ is obtained by cross product of $N_1(s)$ and $N_2(s)$ for each s . The point cloud data D belonging to these longitudinal planes are identified for each s by a minimum distance criterion using dot product. The points satisfying this criterion are further used to obtain r_h where h is the number of points which lie on the longitudinal plane located at s . At each longitudinal plane defined by $(N_1(s), N_2(s), N_3(s))$ a collection of vectors r_h exists in Cartesian coordinates that describe the distance from the center line point $\rho(s)$ to surface points D . To more efficiently analyze the data on a longitudinal plane basis, a transformation to polar coordinates takes place. In polar coordinates the magnitudes of r_h vectors represent radius. By considering a suite of dot products between each r_h vector, $N_1(s)$, and $N_2(s)$ within each longitudinal plane at a given s the angular values θ within that plane associated with each r_h is obtained. This procedure is summarized in Algorithm 2.

Algorithm 2 Surface Parameterization

Input:

$$D = [d_\alpha(1), d_\alpha(2), \dots, d_\alpha(n)]^T \quad \forall \alpha \in \{x, y, z\}$$

▷ Center line of Aorta

$$\rho(s)$$

Output:

$$N(s) = \text{col}(N_1(s), N_2(s), N_3(s))$$

▷ Longitudinal Planes along s

$$r_h(s, \theta)$$

Algorithm:

for all s **do**

$$N_1(s) \leftarrow \frac{\partial \rho(s)}{\partial s}$$

$$N_2 \leftarrow \frac{(e_x - N_1 \cdot e_x)}{\|e_x - N_1 \cdot e_x\|}$$

$$N_3(s) \leftarrow N_1(s) \times N_2(s)$$

end for

for all s **do** $h \leftarrow 0$

for all $i \in \mathcal{R}^n$ **do** $h \leftarrow h + 1$

if $|N_1(s) \cdot v_{\rho(s)/D(i)}| < 0.01$ **then**

$$r_h(s) \leftarrow v_{\rho(s)/D(i)}$$

end if

end for

$$\text{GETANGLE}(\mathbf{N}, r_h(s))$$

end for

function $\text{GETANGLE}(\mathbf{N}, r_h(s))$

if $r_h(s) \cdot N_2(s) = 1$ **then**

$$\theta \leftarrow 0$$

else if $r_h(s) \cdot N_2(s) = -1$ **then**

$$\theta \leftarrow \pi$$

else if $(|r_h(s) \cdot N_2(s)| \geq 0 \text{ and } r_h(s) \cdot N_3(s) \geq 0)$ **then**

$$\theta \leftarrow \cos^{-1}(r_h(s) \cdot N_2(s))$$

else $(|r_h(s) \cdot N_2(s)| \geq 0 \text{ and } r_h(s) \cdot N_3(s) < 0)$

$$\theta_h(s) \leftarrow -\cos^{-1}(r_h(s) \cdot N_2(s)) + 2\pi$$

end if

end function

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] C. M. Porth, *Essentials of pathophysiology: Concepts of altered health states*. Lippincott Williams & Wilkins, 2010.
- [2] A. R. Zankl, H. Schumacher, U. Krumsdorf, H. A. Katus, L. Jahn *et al.*, “Pathology, natural history and treatment of abdominal aortic aneurysms,” *Clinical Research in Cardiology*, vol. 96, no. 3, pp. 140–151, 2007.
- [3] A. Klink, F. Hyafil, J. Rudd, P. Faries, V. Fuster, Z. Mallat, O. Meilhac, W. J. Mulder, J.-B. Michel, F. Ramirez *et al.*, “Diagnostic and therapeutic strategies for small abdominal aortic aneurysms,” *Nature Reviews Cardiology*, vol. 8, no. 6, pp. 338–347, 2011.
- [4] H. Kniemeyer, T. Kessler, P. U. Reber, H. B. Ris, H. Hakki, and M. K. Widmer, “Treatment of ruptured abdominal aortic aneurysm, a permanent challenge or a waste of resources? prediction of outcome using a multi-organ-dysfunction score,” *European Journal of Vascular and Endovascular Surgery*, pp. 190–196, 2000.
- [5] M. Wassef, B. T. Baxter, R. L. Chisholm, R. L. Dalman, M. F. Fillinger, J. Heinecke, J. D. Humphrey, H. Kuivaniemi, W. C. Parks, W. H. Pearce *et al.*, “Pathogenesis of abdominal aortic aneurysms: a multidisciplinary research program supported by the national heart, lung, and blood institute,” *Journal of Vascular Surgery*, vol. 34, no. 4, pp. 730–738, 2001.
- [6] S. Einav, J. Ricotta, and D. Bluestein, “Abdominal aortic aneurysm risk of rupture: patient-specific FSI simulations using anisotropic model,” *Journal of Biomechanical Engineering*, vol. 131, pp. 031 001–1, 2009.
- [7] M. F. Fillinger, M. L Raghavan, S. P. Marra, J. L. Cronenwett, F. E. Kennedy *et al.*, “In vivo analysis of mechanical wall stress and abdominal aortic aneurysm rupture risk.” *Journal of Vascular Surgery*, vol. 36, 2002.
- [8] D. Bluestein, K. Dumont, M. De Beule, J. Ricotta, P. Impellizzeri, B. Verheghe, and P. Verdonck, “Intraluminal thrombus and risk of rupture in patient specific abdominal aortic aneurysm–FSI modelling,” *Computer Methods in Biomechanics and Biomedical engineering*, vol. 12, no. 1, pp. 73–81, 2009.

- [9] B. Wolters, M. C. M. Rutten, G. W. H. Schurink, U. Kose, J. De Hart, and F. N. Van De Vosse, "A patient-specific computational model of fluid–structure interaction in abdominal aortic aneurysms," *Medical Engineering & physics*, pp. 871–883, 2005.
- [10] F. H. Epstein, G. H. Gibbons, and V. J. Dzau, "The emerging concept of vascular remodeling," *New England Journal of Medicine*, vol. 330, no. 20, pp. 1431–1438, 1994.
- [11] J. D. Humphrey, *Cardiovascular solid mechanics: cells, tissues, and organs*. Springer Verlag, 2002.
- [12] N. Resnick, H. Yahav, A. Shay-Salit, M. Shushy, S. Schubert, L. C. M. Zilberman, and E. Wofovitz, "Fluid shear stress and the vascular endothelium: for better and for worse," *Progress in Biophysics and molecular biology*, vol. 81, no. 3, pp. 177–199, 2003.
- [13] A. B. Driss, J. Benessiano, P. Poitevin, B. I. Levy, and J.-B. Michel, "Arterial expansive remodeling induced by high flow rates," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 272, no. 2, pp. H851–H858, 1997.
- [14] M. Zamir, "Shear forces and blood vessel radii in the cardiovascular system." *The Journal of General physiology*, vol. 69, no. 4, pp. 449–461, 1977.
- [15] M. A. Hajdu and G. L. Baumbach, "Mechanics of large and small cerebral arteries in chronic hypertension," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 266, no. 3, pp. H1027–H1033, 1994.
- [16] J.-J. Hu, S. Baek, and J. D. Humphrey, "Stress–strain behavior of the passive basilar artery in normotension and hypertension," *Journal of biomechanics*, vol. 40, no. 11, pp. 2559–2563, 2007.
- [17] J. D. Humphrey, J. Eberth, W. Dye, and R. L. Gleason, "Fundamental role of axial stress in compensatory adaptations by arteries," *Journal of biomechanics*, vol. 42, no. 1, pp. 1–8, 2009.
- [18] Z. S. Jackson, D. Dajnowiec, A. I. Gotlieb, and B. L. Langille, "Partial off-loading of longitudinal tension induces arterial tortuosity," *Arteriosclerosis, thrombosis, and vascular biology*, vol. 25, no. 5, pp. 957–962, 2005.
- [19] Z. S. Jackson, A. I. Gotlieb, and B. L. Langille, "Wall tissue remodeling regulates longitudinal tension in arteries," *Circulation research*, vol. 90, no. 8, pp. 918–925, 2002.

- [20] S. Baek, K. R. Rajagopal, J. D. Humphrey *et al.*, “A theoretical model of enlarging intracranial fusiform aneurysms,” *TRANSACTIONS-ASME JOURNAL OF BIOMECHANICAL ENGINEERING*, vol. 128, no. 1, p. 142, 2006.
- [21] S. Baek, A. Valentin, and J. D. Humphrey, “Biochemomechanics of cerebral vasospasm and its resolution: II. constitutive relations and model simulations,” *Annals of biomedical engineering*, vol. 35, no. 9, pp. 1498–1509, 2007.
- [22] R. L. Gleason and J. D. Humphrey, “A mixture model of arterial growth and remodeling in hypertension: altered muscle tone and tissue turnover,” *Journal of vascular research*, vol. 41, no. 4, pp. 352–363, 2004.
- [23] —, “Effects of a sustained extension on arterial growth and remodeling: a theoretical study,” *Journal of biomechanics*, vol. 38, no. 6, pp. 1255–1261, 2005.
- [24] S. Zeinali-Davarani, A. Sheidaei, and S. Baek, “A finite element model of stress-mediated vascular adaptation: application to abdominal aortic aneurysms,” *Computer methods in biomechanics and biomedical engineering*, vol. 14, no. 9, pp. 803–817, 2011.
- [25] A. Sheidaei, S. C. Hunley, S. Zeinali-Davarani, L. G. Raguin, and S. Baek, “Simulation of abdominal aortic aneurysm growth with updating hemodynamic loads using a realistic geometry,” *Medical engineering & physics*, vol. 33, no. 1, pp. 80–88, 2011.
- [26] S. Zeinali-Davarani, J. Choi, and S. Baek, “On parameter estimation for biaxial mechanical behavior of arteries,” *Journal of biomechanics*, vol. 42, no. 4, pp. 524–530, 2009.
- [27] J. S. Wilson, S. Baek, and J. D. Humphrey, “Parametric study of effects of collagen turnover on the natural history of abdominal aortic aneurysms,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, vol. 469, no. 2150, 2013.
- [28] S. Zeinali-Davarani, L. G. Raguin, D. A. Vorp, and S. Baek, “Identification of in vivo material and geometric parameters of a human aorta: toward patient-specific modeling of abdominal aortic aneurysm,” *Biomechanics and modeling in mechanobiology*, vol. 10, no. 5, pp. 689–699, 2011.
- [29] J. Kim and S. Baek, “Circumferential variations of mechanical behavior of the porcine thoracic aorta during the inflation test,” *Journal of biomechanics*, vol. 44, no. 10, pp. 1941–1947, 2011.

- [30] S. T. Kwon, J. E. Rectenwald, S. Baek *et al.*, “Intrasac pressure changes and vascular remodeling after endovascular repair of abdominal aortic aneurysms: review and biomechanical model simulation.” *Journal of biomechanical engineering*, vol. 133, no. 1, p. 011011, 2011.
- [31] Y. Xu, J. Choi, and S. Oh, “Mobile sensor network navigation using gaussian processes with truncated observations,” *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1118–1131, 2011.
- [32] Y. Xu and J. Choi, “Adaptive sampling for learning Gaussian processes using mobile sensor networks,” *Sensors*, vol. 11, no. 3, pp. 3051–3066, 2011.
- [33] A. Ijaz, J. Choi, W. Lee, and S. Baek, “Prediction of abdominal aortic aneurysms using sparse gaussian process regression,” *Proceedings of the ASME 2013 Summer Bioengineering Conference*, 2013.
- [34] C. Williams and C. Rasmussen, “Gaussian processes for machine learning,” *MIT Press*, 2006.
- [35] A. J. Smola and P. Bartlett, “Sparse greedy Gaussian process regression,” in *Advances in Neural Information Processing Systems 13*. Citeseer, 2001.
- [36] C. Williams and M. Seeger, “Using the nyström method to speed up kernel machines,” in *Advances in Neural Information Processing Systems 13*. Citeseer, 2001.
- [37] N. D. Lawrence, M. Seeger, and R. Herbrich, “Fast sparse Gaussian process methods: The informative vector machine,” *Advances in neural information processing systems*, vol. 15, no. 15, pp. 609–616, 2002.
- [38] M. Seeger, “Bayesian Gaussian process models: Pac-bayesian generalisation error bounds and sparse approximations,” 2003.
- [39] V. Tresp, “A bayesian committee machine,” *Neural Computation*, vol. 12, no. 11, pp. 2719–2741, 2000.
- [40] A. Brix and P. J. Diggle, “Spatiotemporal prediction for log-Gaussian cox processes,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 4, pp. 823–841, 2001.
- [41] A. J. Storkey, “Truncated covariance matrices and toeplitz methods in Gaussian processes,” in *Artificial Neural Networks*, vol. 1. IET, 1999, pp. 55–60.

- [42] S. S. Raut, S. Chandra, J. Shum, and E. A. Finol, "The role of geometric and biomechanical factors in abdominal aortic aneurysm rupture risk assessment," *Annals of biomedical engineering*, pp. 1–19, 2013.
- [43] M. Gaudard, M. Karson, E. Linder, and D. Sinha, "Bayesian spatial prediction," *Environmental and Ecological Statistics*, vol. 6, no. 2, pp. 147–171, 1999.
- [44] Y. Xu, J. Choi, S. Dass, and T. Maiti, "Sequential bayesian prediction and adaptive sampling algorithms for mobile sensor networks," *IEEE Transactions on Automatic Control*, vol. 57, no. 8, pp. 2078–2084, 2012.
- [45] J. Weng and B. Lee, "Event detection in twitter," *Proc. of ICWSM*, 2011.
- [46] M. Paul and M. Dredze, "You are what you tweet: Analyzing twitter for public health," in *Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011)*, 2011.
- [47] A. Tumasjan, T. Sprenger, P. Sandner, and I. Welp, "Predicting elections with twitter: What 140 characters reveal about political sentiment," in *Proceedings of the fourth international aai conference on weblogs and social media*, 2010, pp. 178–185.
- [48] S. Verma, S. Vieweg, W. Corvey, L. Palen, J. Martin, M. Palmer, A. Schram, and K. Anderson, "Natural language processing to the rescue?: Extracting'situational awareness' tweets during mass emergency," *Proc. ICWSM*, 2011.
- [49] B. OConnor, R. Balasubramanyan, B. Routledge, and N. Smith, "From tweets to polls: Linking text sentiment to public opinion time series," in *Proceedings of the International AAAI Conference on Weblogs and Social Media*, 2010, pp. 122–129.
- [50] S. Morinaga, K. Yamanishi, K. Tateishi, and T. Fukushima, "Mining product reputations on the web," in *Conference on Knowledge Discovery in Data: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, vol. 23. Citeseer, 2002, pp. 341–349.
- [51] J. Blitzer, M. Dredze, and F. Pereira, "Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification," in *Annual Meeting-Association For Computational Linguistics*, vol. 45, no. 1, 2007, p. 440.
- [52] R. Balasubramanyan, W. Cohen, D. Pierce, and D. Redlawsk, "Modeling polarizing topics: When do different political communities respond differently to the same news?" in *Sixth International AAAI Conference on Weblogs and Social Media*, 2012.

- [53] L. Xiang, Q. Yuan, S. Zhao, L. Chen, X. Zhang, Q. Yang, and J. Sun, “Temporal recommendation on graphs via long-and short-term preference fusion,” in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2010, pp. 723–732.
- [54] B. Pang and L. Lee, *Opinion mining and sentiment analysis*. Now Pub, 2008.
- [55] M. De Choudhury, Y. Lin, H. Sundaram, K. Candan, L. Xie, and A. Kelliher, “How does the data sampling strategy impact the discovery of information diffusion in social media,” in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, 2010, pp. 34–41.
- [56] B. Pang and L. Lee, “A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts,” in *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 2004, p. 271.
- [57] J. Wiebe, R. Bruce, and T. O’Hara, “Development and use of a gold-standard data set for subjectivity classifications,” in *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*. Association for Computational Linguistics, 1999, pp. 246–253.
- [58] Y. Mejova and P. Srinivasan, “exploring feature definition and selection for sentiment classifiers,” in *Proceedings of the Fifth international aaai conference on Weblogs and Social media (icWSm-2011)*, 2011.
- [59] W. Peng and D. Park, “generate adjective sentiment dictionary for social media sentiment analysis using constrained nonnegative matrix factorization,” *Urbana*, vol. 51, p. 61801, 2004.
- [60] Y. Lu, M. Castellanos, U. Dayal, and C. Zhai, “Automatic construction of a context-aware sentiment lexicon: an optimization approach,” in *Proceedings of the 20th international conference on World wide web*. ACM, 2011, pp. 347–356.
- [61] S. Baccianella, A. Esuli, and F. Sebastiani, “Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining,” in *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC10)*, Valletta, Malta, May. European Language Resources Association (ELRA), 2010.
- [62] A. Raftafi, V. Guigue, and P. Gallinari, “Coping with the document frequency bias in sentiment classification,” in *Sixth International AAAI Conference on Weblogs and Social Media*, 2012.

- [63] L. Chen, W. Wang, M. Nagarajan, S. Wang, and A. Sheth, “Extracting diverse sentiment expressions with target-dependent polarity from twitter,” in *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2012, pp. 50–57.
- [64] D. Davidov, O. Tsur, and A. Rappoport, “Enhanced sentiment learning using twitter hashtags and smileys,” in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*. Association for Computational Linguistics, 2010, pp. 241–249.
- [65] E. Kouloumpis, T. Wilson, and J. Moore, “Twitter sentiment analysis: The good the bad and the omg,” in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 2011.
- [66] E. Gilbert and K. Karahalios, “Widespread worry and the stock market,” in *Proceedings of the international conference on weblogs and social media*, vol. 2, no. 1, 2010, pp. 229–247.
- [67] M. Koppel and I. Shtrimberg, “Good news or bad news? let the market decide,” *Computing attitude and affect in text: Theory and applications*, pp. 297–301, 2006.
- [68] P. S. Dodds and C. M. Danforth, “Measuring the happiness of large-scale written expression: Songs, blogs, and presidents,” *Journal of Happiness Studies*, vol. 11, no. 4, pp. 441–456, 2010.
- [69] N. Nori, D. Bollegala, and M. Ishizuka, “Exploiting user interest on social media for aggregating diverse data and predicting interest,” in *Fifth International AAAI Conference on Weblogs and Social Media*, 2011.
- [70] O. Dabeer, P. Mehendale, A. Karnik, and A. Saroop, “Timing tweets to increase effectiveness of information campaigns,” *Proc. ICWSM*, 2011.
- [71] M. Naaman, A. Zhang, S. Brody, and G. Lotan, “On the study of diurnal urban routines on twitter,” in *Sixth International AAAI Conference on Weblogs and Social Media*, 2012.
- [72] C. Manning and H. Schütze, *Foundations of statistical natural language processing*. MIT press, 1999.
- [73] D. B. Peizer and J. W. Pratt, “A normal approximation for binomial, f, beta, and other common, related tail probabilities, i,” *Journal of the American Statistical Association*, vol. 63, no. 324, pp. 1416–1456, 1968.

- [74] “Twitter votes to obama.” [Online]. Available:
<http://www.buzzfeed.com/jwherrman/twitter-users-say-they-voted-for-obama-2-to-1>
- [75] A. Go, R. Bhayani, and L. Huang, “Twitter sentiment classification using distant supervision,” *CS224N Project Report, Stanford*, pp. 1–12, 2009.
- [76] A. Pak and P. Paroubek, “Twitter as a corpus for sentiment analysis and opinion mining,” in *Proceedings of LREC*, vol. 2010, 2010.
- [77] C. Fellbaum, “Wordnet,” *Theory and Applications of Ontology: Computer Applications*, pp. 231–243, 2010.