Project Plan

Gareth Lomax

June 2019

1 Supervisor Details

1.1 Project Supervisor (External)

Dr Nils W. Metternich
Associate Professor in International Relations
School of Public Policy
Dept Of Political Science UCL
Email: n.metternich@ucl.ac.uk

1.2 Supervisors (Internal)

Prof Sanjeev Gupta
Professor of Earth Science
Faculty of Engineering, Department of Earth Science Engineering
Imperial College London
Email: s.gupta@imperial.ac.uk

Dr Steven Banham
Research Associate
Faculty of Engineering, Department of Earth Science Engineering
Imperial College London
Email: s.banham@imperial.ac.uk

2 Introduction and Motivation

Civil, interstate and terrorist conflict holds a massive influence in dictating the future of a country. The ability to forewarn conflict events and identify areas at high risk of conflict ignition allows the international community to take action to prevent the carnage and economic disruption that accompanies war. While a number of conflict prediction systems exist, such as the ICEWS project run by the US Dept of Defense O'Brien [2010], few large scale conflict forecasting efforts that openly broadcast their predictions to the public currently exist Hegre et al. [2019a]. As policy decisions are increasingly made based off model predictions, and given the poor performance of expert judgement Tetlock [2018] it is vital that a robust range of conflict prediction tools are produced in order to assist NGOs and at risk groups with strategic decisions, with the eventual goal of pre-empting and reducing the damage caused by potential conflicts though an early warning system. In practice conflict predictions usually consists of a binary variable for each spatio-temporal unit of analysis denoting whether one or more battle related deaths can be attributed to a violent conflict in a specified region in space and time. Current efforts in conflict prediction are centered around theoretical models, agent based modelling and statistical machine learning approaches. The most promising efforts have used ensembles of machine learning approaches based on large dissagregated data sets Hegre et al. [2019a]. Primarily the machine learning approaches have centered around variants of time lagged logistic regression and random forests Muchlinski et al. [2016], due to their effectiveness in a wide range of problems, and due the inherent class imbalance present in conflict data. This project seeks to explore the efficacy of deeper machine learning techniques when applied to conflict data. Convolutional Long Short-Term Memory (convLSTM) Shi et al. [2015] models have produced best in class results in a number of spatio-temporal prediction tasks such as weather nowcasting Shi et al. [2015]. We aim to explore whether the advantages of using techniques which explicitly rely on spatially encoded information can more effectively predict spatial variance in conflicts. The main aim of this project is to asses the suitability of convolutional LSTM to predict conflict. This will be done by comparing the ability of an implemented convLSTM model to predict the presence of conflict in each (55×55)km cell a vector grid representation of Africa over a 12 month forecasting period. In order to achieve this we define four subsidiary aims are to be completed: (**i**) To implement and verify a convLSTM model in python using the Pytorch architecture. (**ii**) To adapt conflict predictor data taken from the PRIO and UCDP data projects (see section 3.1.5) into a subsampled 5th dimensional tensor structure to provide a dataset suitable for both this project and future projects seeking to use convolutional predictive techniques. (**iii**) To use the implemented convLSTM module as a building block for a deep conflict prediction model, training on the prepared dataset. (**iiii**) To compare the predictive ability of the produced model with existing approaches Hegre et al. [2017], in order to asses the efficacy of the proposed method.

3 Literature Review

3.1 Conflict

3.1.1 Terminology

Note that from here we use the UCDP definition of conflict which defines a conflict as a contested incompatibility between two parties which results in at least 25 battle related deaths per year Croicu and Sundberg [2016]. We will also discuss various data that have been identified as having relevant to conflict prediction. This data is usually identified as relevant from theoretical or agent based studies targeting specific conflict causation systems e.g imbalances in wealth distribution between different ethnic groups in a region Mitra and Ray [2014]. Once identified in the literature these data are then typically used as input data into machine learning algorithms in subsequent studies. In line with the machine learning community we will refer to these data as predictors. Predictors usually describe the physical and social geography of a region in space, as well as the economic climate Hegre and Sambanis [2006]. They may be categorical data, such as the binary presence of oil wells in the region they describe, or continuous, such as the nighttime luminescence of a region, as measured by satellite Tollefsen et al. [2012].

3.1.2 Overview

Broadly we may separate conflict prediction efforts into two categories; Model driven and Data driven. Model driven approaches seek to impose theoretical or semi-theoretical rule based frameworks to predict future conflict usually in the form of agent based modelling, whereas data drive approaches leverage machine learning methods, such as random forests or logistic regression Muchlinski et al. [2016]. Machine learning approaches are advantageous when seeking overall predictive power, however they often lack transparency, and are less successful when modelling factors surrounding specific causes of conflict. Hegre et al. [2017]It should be noted however that many of the apriori theoretical models are still informed by observation of the data by subject experts. Theoretical and agent based models usually center about investigating specific conflict causation systems, and often much of the choice of predictors in data driven efforts is based on results from strict agent based models which have highlighted the roles of specific conflict predictors, such as the exclusion of different ethnic groups or climate change Sarsons [2015] Weidmann [2011]. Models use a variety of spatial representations to predict conflict. Originally predictions focused on large scale units of analysis, using a country by country representation, using country wide aggregated predictor variables, such as GDP, to predict the presence of conflict in the country as a whole. Numerous studies have noted the advantage of disagregating data into smaller geographic representations, such as Weidmann and Ward [2010], which represents Bosnia using sub-national municipalities, and forecasts the probability of violence in each municipality, or Hegre et al. [2019a] which uses a continent wide Cartesian grid representation to forecast conflict on a cell grid level. Other works such as Guo et al. [2016] have represented countries as city networks to predict conflict due to the ease of representing both geographic and political connections using single or multi layer networks. They use predictor variables corresponding to city entities rather than predictors dependent on geographic regions.

3.1.3 Conflict Predictors

A wide array of relevant systems have been identified as contributing to conflict. Collier and Hoeffler [2000] identifies low economic growth as a significant cause of conflict due to its reductive effect on a states ability to stamp out protests and violence. This conclusion is furthered in Fearon and Laitin [2003], in which they authors demonstrate that poor economic progress raises the opportunity cost involved with rebellion and contesting civil wars and that poor economic conditions favour rebel recruitment. Mitra and Ray [2014] Uses econometric theory applied to Hindu-Muslim violence in India and examines the effect of raising incomes of either of the groups, concluding that an income shock is likely to lead to grievance and violence between the groups. Alongside this Olsson [2007] finds that the abundance of precious resources typically leads to lower economic progress and an increased susceptibility to conflict. Rainfall variation is often used as a proxy to study the effect of income shocks on conflict. Bohlken and Sergenti [2010] have used this to find correlation between the prevalence of rioting and rainfall. Sarsons [2015] expands on this and finds that rainfall still acts well as a predictor of unrest even in areas adjacent to dams, which should be immune to sudden droughts, suggesting that climate has a larger role to play. In Hegre and Sambanis [2006] the authors conduct sensitivity analyses on 88 variables commonly used in conflict prediction with the objective of giving a reality check in a field containing a large amount of contradictory findings. They again support the idea of the strong effect of GDP on conflict likelihood, demonstrating that a 1 standard deviation reduction of income across a region raises the probability of civil war by 65 percent. They also show that regions containing mountainous and rough terrain can also predict the presence of violence well, in agreement with Fearon and Laitin [2003], a study which reasons that the natural hideouts favour insurgent groups hiding there. Alongside this they also demonstrate that large militaries prevent intrastate conflict and recent political instability is the most robust predictor for civil war. The presence of civil war in adjacent countries is also noted as a robust predictor, as would be expected given the large range of work on conflict diffusion and contagion.

3.1.4 Spatial Dependence

Clear amongst each of these works is the importance of the relative position of competing groups in conflict prediction. The effect of position of conflict actors is of huge importance in conflict prediction, both due to many conflict predictors being a function of location (avg annual temperature, presence of oil resources, susceptibility to climate change ect) and which are often clustered**?**, but also due to the phenomena of conflict contagion and diffusion. Weidmann and Ward [2010] Compares two models of conflict applied to geolocated conflict data from the Bosnian war (March 1992 - October 1995) and demonstrates the improvement in out of sample prediction of violence prediction (binary: presence vs absence) of 3% accuracy by including nearest neighbours spatial information and demonstrates the tendency of conflict to spill over between neighbouring groups. Metternich et al. [2017] Furthers this line of enquiry and demonstrates contagion between different conflict actors as being reliant on both of the sending and receiving groups utilising an bi-linear network model, and demonstrate that the effect that the ethnic exclusion of both the sending and receiving conflict groups is dependant on the geographic proximity of the two groups. Guo et al. [2016] Use a geographic city network based model and show that the betweenness centrality of a city can explain above 80% of the variance in terrorist attacks in city location, and use an agent based model to demonstrate this is as a result of the high likelihood of fuzzy cultural boundaries at high betweeness centrality locations.

3.1.5 Data

The project will use event and predictor data from the PRIO grid and UCDP-GED data sets Tollefsen et al. [2012] Sundberg and Melander [2013] . The PRIO (Peace Research Institute of Oslo) data set is widely used in conflict research to provide predictor information at a sub national level. The PRIO grid consists of a vector cell grid of dimensions (0.5 x 0.5) degrees of latitude and longitude covering the globe. Each cell contains variables sorted into conflict relevant groups (socioeconomic, resources and climate.) The UCDP data set (Uppsala conflict data program) encodes the number of deaths due to a conflict on a monthly basis. For a given political situation to be classed as a conflict, it must meet a threshold of 25 battle related fatalities per year. For a conflict to be recorded all groups involved in the conflict must be clearly defined. Croicu and Sundberg [2016]. A version of the UCDP base has been adapted to fit the PRIO grid structure, and covers the years 1989-2018. The UCDP encodes the number of deaths at a given grid cell at each month. The norm in conflcit prediction is to reduce this to a binary presence variable, where efforts are focused on the prediction of presence of violence, rather than its magnitude. We will shape our prediction efforts around this for ease of comparison with other efforts. We note that due to the complexities involved with extracting data from conflict zones, the data used will require spatial imputation in approximately 5% of cases.

3.1.6 State of The Art

The Views Project Hegre et al. [2019a] Hegre et al. [2019b] provides a clear example of the state of the art in conflict prediction. The Views Project relies on the PRIO GRID data structure, and is currently the leading publicly available real-time conflict forecasting model. Comparable approaches have previously been developed by the US Dept of Defense and the CIA, but these approaches are not readily accessible. Views combines an ensemble of models based on thematic aggregations of predictors (Baseline, Conflict History, Natural geography, social geography, country level variables and protest) to predict for 36 future months on two scales: on a country by country bases and PRIO grid cell level. They predict the presence of violence (binary) comparing with the UCDP GED database (version 18.1) using dynamic simulation and one step ahead variations of logit and random forest models. see Muchlinski et al. [2016] for a comparison of random forest and logit.
Views has produced several results of note that are worth considering. Primarily they demonstrate that dissagregated gird data outperforms country level data in terms of prediction accuracy (84% vs 99.1% for a full ensemble run) due to the advantages that sub-national dissagregation has when modeling conflict spill over. They also compare the inclusion of a wide number of variables and find the greatest predictive benefit from the inclusion of natural and social geographic features.

Their work also highlights a number of pitfalls with testing and validation which should be considered when applying machine learning techniques to conflict. They isolate the months January 2015 - December 2017 for use as a test set, however they also note the static nature of conflict in Africa in this time frame, implying that their test set may be comparatively "easy", and their accuracy inflated. Even if slightly inflated the accuracies and ROC curves presented demonstrate an incredibly impressive quality of prediction of conflict. Their model both predicts accurately and represents conflict contagion well. We should also note that their use of standardised evaluations of their prediction quality (AUROC, AUPR, Brier score and Accuracy) and pipeline approach allow for easy comparison and bench marking of future conflict prediction efforts. The Views team continue to publish predictions every month which can be found at: https://www.pcr.uu.se/research/views/current-forecasts/

3.2 Sequence Prediction

Conflict prediction, in essence, is a spatio-temporal sequence prediction problem. As has been noted above in section 3.1.3, conflict is highly dependant on spatial configurations. As a result we find it prudent to seek the use of methods that are reliant on spatial encoding and have produced outstanding results in spatio-temporal prediction tasks.

3.2.1 Recurrent Neural Nets

Recurrent Neural Nets (RNNs) and their variants are the preferred machine learning architecture for temporal problems. The RNN is an extension of a standard feed forward neural net designed to handle temporally related sequences of input data, which may be of variable length. Graves [2013] This is achieved by the introduction of a recurrent hidden layer h_t whose activation is a weighted combination of the previous hidden layer h_{t-1} and the input at the current time step x_t. This differentiates RNNs from feed forward neural networks: the hidden layer h can be said to be self connected. Graves [2012] This may be seen visualised in figure 1. Note that it is common practice to represent the RNN in an "unfolded state" to visualise the output process, and is especially useful in deep RNN networks. RNNs take a sequence of inputs $\mathbf{x}$ and map to a sequence of outputs $\mathbf{y}$:

$$
\begin{aligned}
h_t &= \mathcal{H}(W_{ih}x_t + W_{hh}h_{t-1} + b_h) \\
y_t &= W_{hy}h_t + b_y
\end{aligned}
\tag{1}
$$

Where h_t is the hidden layer, y_t is the output, x_t is the input, W_{ih}, W_{hh} and W_{hy} are the weight matrices for the fully connected layers and b_h and b_y are biases. $\mathcal{H}$ is the hidden layer activation function, which in practice is usually a sigmoid. Graves [2012] Graves [2013] RNNs perform well on sequence based tasks due to residual information from preceding inputs remaining in the network. In simplistic terms; they remember the previous inputs and may draw inferences based on the sequence as a whole. Unfortunately basic RNNs are unreliable when applied to extracting long temporal relationships in the data due to issues with gradients vanishing and exploding whilst training via back propogation, as has been demonstrated in Hochreiter [2003]. As a result of this flaw training with back propagated gradient descent models is often unfeasible or impossible. Several additions to vanilla RNN architectures have been introduced to counteract this flaw. Aside from efforts to produce training methods immune to the gradient problems. The Long Short-Term Memory architecture (LSTM) was proposed in 1997 in **?**, although the method went relatively unexplored for a number of years, before being used to produce an array of best in class results. In the time between several additions to the LSTM structure have been proposed and accepted by the community. Gated Recurrent Units (GRUs)Chung et al. [2014] were also introduced in 2014 and have been involved in healthy competition with LSTMs with both producing impressive results. Both LSTMs and GRUs introduce gated memory approaches to the activation functions of the RNN hidden layer to overcome the vanishing gradient issue. We choose to focus on LSTMs due to the greater amount one may tailor their activation functions. Here we define the LSTM:

$$
\begin{aligned}
i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \\
f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \\
c_t &= f_t \circ c_{t-1} + i_t \circ \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\
o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + w_{co} \circ c_t + b_o) \\
h_t &= o_t \circ \tanh(c_t)
\end{aligned}
\tag{2}
$$

Where f_t, i_t, $o_t$and c_t are the forget gates, input gates, output gates and cell memories respectively. σ is the sigmoid activation function. In this case $a \circ b$ denotes element wise multiplication. As we can see from the above equations the role of the memory cell is significant. It enables longer term memory storage as the cell memory is selectively updated, instead of updating at every time step as in a vanilla RNN. As a result it retains significant features for longer. In other words, this allows information to skip multiple time steps if the information is relevant. The FC-LSTM is a trivial extension of the LSTM where the x_t and h_t layers are extended to fully connected vectors, allowing large amounts of interconnected information to travel through the LSTM at each time step. This is also useful as it allows categorical information to be used in LSTMs by utilising one-hot encoding Rodríguez et al. [2018]. FC-LSTMs have performed very well in temporally dependant tasks, including in natural language processing (NLP) and sequence prediction. Notable results include Vinyals et al. [2015] in which the state of the art BLEU-1 captioning score was raised from 25 to 59. While FC-LSTMs are highly performant in sequence to sequence (seq2seq) tasks involving suitably linear data such as streams of encoded words, a number of works have explored methods using LSTMs to process spatially related data. Convolutional methods are the clear state of the art when extracting information from gridded data sources, especially image data Sharma et al. [2018]. Approaches such as those documented in Srivastava et al. [2015] use pretrained convolutional neural nets (CNNs) to encode image data at each sequence step, taking representations from either the networks high level percepts or directly from the final fully connected

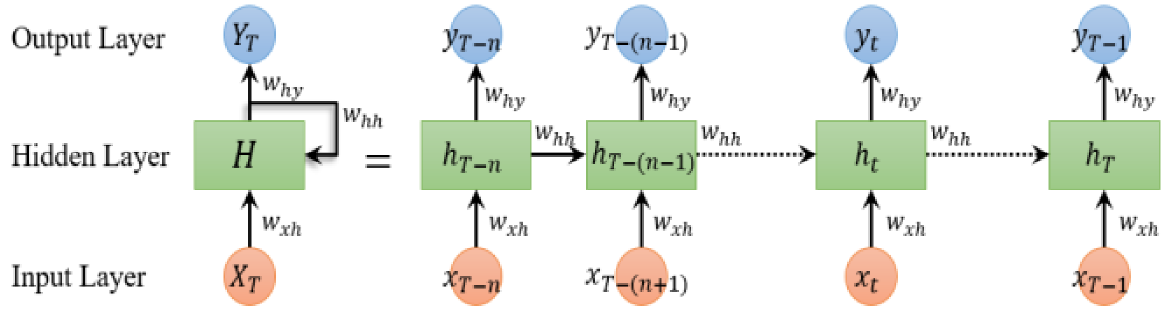

Figure 1: Diagram of folded and unfolded RNN. At timestep t the RNN takes an input vector x_t which combined with the the hidden layer from the previous timestep, h_{t-1} and a bias term, b_h according to an activation function Graves [2013]. The output a given timestep y_t is a weighted fully connected transformation of the hidden state. The incorporation of the fully connected layer, modulated by W_{hh} allows sequence information pass between different sequence steps, allowing the RNN to recognise patterns present in the data. The RNN is represented in its "unfolded" format on the right, demonstrating how the hidden state passes through the system as time progresses. Image source: Cui et al. [2018]

layer and passing this to the FC-LSTM. This approach has been explored in both supervised and unsupervised fashions to produce good results in learning video representations. They also explore the use of LSTM for future prediction using an implementation of the encoder-decoder system first proposed in Sutskever et al. [2014]. While this works well at the input to state stage as the spatial relationships are captured by the CNN input into the LSTM, the FC-LSTM does not preserve the spatial encoding in the state to state transitions due to its fully connected nature. I.e once the spatial data has been extracted from the image when entering the LSTM, it may interact with other information free of spatial constraints at each time step. Shi et al. [2015] proposes a further advancement of the LSTM, and implements a fully convolutional LSTM (convLSTM). Where the state to state transitions alongside the input to state transitions are all convolutional. The governing equations from 2 are replaced with their convolutional alternatives to give:

$$\begin{aligned}
i_t &= \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \\
f_t &= \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f) \\
C_t &= f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \\
o_t &= \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + w_{co} C n_t + b_o) \\
H_t &= o_t \tanh n(C_t)
\end{aligned} \tag{3}$$

Here C_t, H_t and X_t are the tensor equivalents of c_t, h_t and x_t in 2, and $*$ denotes a convolution operation. The state to state convolutions in the convLSTM act to preserve the spatial structure of the data, in essence limiting the rate at which information can traverse the system at each timestep. A larger convolutional kernel captures higher speed spatial effects. A 1x1 kernel in effect acts as a FC-LSTM. The convLSTM has been shown to outperform FC-LSTM in spatio-temporal tasks, as demonstrated in Hong et al. [2017] Gehring et al. [2017]. The structure of the convLSTM is appealing to those seeking to predict conflict, due to its capability to combine input from a deep array of predictor variable layers whilst respecting the spatial relationships between the neighbor sites.

4 Proposed Approach

Our proposed approach is to use a convLSTM encoder-decoder model implemented in Pytorch, to predict the presence of conflict as defined by the UCDP GED data project in each PRIO grid cell covering Africa each month over a 12 month period.

4.1 Implementation Details

The code implementation will be primarily written in python using Pytorch, a python implementation of the Torch machine learning library. Pytorch is open source and has become the machine learning library of choice in academia due to its flexibility and relative power. Pytorch does not currently contain an convLSTM implementation. One is present in Keras, however Pytorch is preferred due to its increased performance and flexibility.

4.2 Approach Justification

The choice of a convLSTM model is justified as it has demonstrated best in class results when applied to spatio-temporal prediction tasks, as discussed in 3.1. Conflict has been shown to have strong spatio-temporal dependency, due to the processes

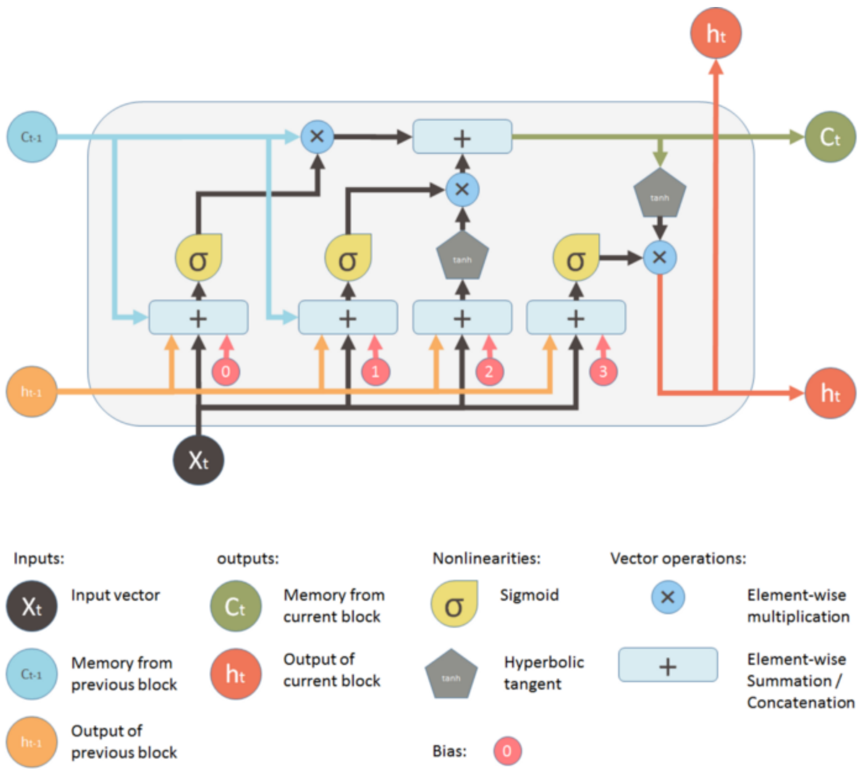

Figure 2: Visual representation of a single LSTM cell. Compare the unfolded contents of the above figure with 1. In the vanilla RNN the hidden layer is updated according to the equations shown in 1. The LSTM replaces this simple constant update at every time step with a logic gate activation method. Note the addition of a persistent cell memory C_t which is selectively updated according to the contents of the forget gate f_t. The selective updating of the cell memory allows the LSTM to retain significant information from the sequence input for longer than a traditional RNN. An LSTM network functions similarly to a RNN network as shown in 1 but with the central cell replaced with a LSTM cell as shown here. Und

of conflict diffusion and contagion, as discussed in section 3.1.4. Current efforts as discussed in sections 3.1 and 3.1.6 use random forest and logistic regression methods, which do not respect the spatial flow of information. As a result we expect the convLSTM to perform well in comparison to these methods, while using a reduced parameter space. We use the UCDP GED data base as our target data to predict, due to the high level of vetting it is subjected to, as well as its wide use across conflict prediction efforts. The PRIO data grid is chosen due to its wide range of predictor variables across a number of themes (see section 3.1.5), and due to its Cartesian grid structure which is necessary for convolutional methods. The use of UCDP and PRIO data also allows us to compare with other predictive works Hegre et al. [2019a] in order to asses the efficacy of our approach, one of the stated aims of this project. Finally we limit the area of prediction to Africa, as it is the continent that corresponds to the most complete data set of conflict events.

4.3 Requirements

We now present the key steps which must be completed to achieve our implementation, noting steps where particular difficulty may lie.

1. Implementation of a convLSTM module using Pytorch. This process contains three sub steps: (**A**) The implementation of the module. (**B**) The implementation of the training and validation functions. (**C**) Validation of the produced code against the MovingMNIST data set, to ensure our implementation works as required.

2. Production of a pipeline to transform the PRIO and UCDP datasets into sub sampled image formats ready for input into the convLSTM model. The output will a 5th dimensional input tensor of dimensions: (b, s, l, SS_w, SS_h) where b is the batch size, s is the number of steps in the training sequence (i.e the number of preceding months our model will use to predict the next moths conflict), l is the number of different predictor variables we supply to our model, and SS_w and SS_h are the dimensions of the subsampled input image.

3. Construction of the model architectures using the convLSTM module implemented in step (1). The performance of the model architectures on the dataset constructed in step (2) will be compared to find the optimum configuration. Following a finalised architecture we will then conduct hyper parameter tuning to improve model performance.

4. We will then explore data augmentation to improve the robustness of the final model. Special care will be taken with augmentation to respect the norms of social data.

5. Comparison of our predicted results with other predictions in the field, in order to asses the validity of the use of convLSTMs to produce conflict, the stated goal our our project.

4.4 Timeline

Completed	Review of current theory
Completed	Requirements analysis
Completed	high level design
7/07/2019	Finish implementation of convLSTM and CNNLSTM
	Verification of convLSTM on MovingMNIST
14/07/2019	First prototype delivered
	Impute Datasets
20/07/2019	Assemble datasets into described structure
25/07/2019	Train Encoder-Decoder Architectures Using Verified convLSTM Architecture
1/08/2019	Outline Report Structure
5/08/2019	Compare various Encoder-Decoder Architectures and consider ensemble methods
10/08/2019	Optimise Hyper Parameters
12/08/2019	Consider Data Augmentation and GANs Implementation
	Implement Pep-8
20/08/2019	Final Draft of Report
30/08/2019	Report Due
30/08/2019	Code and finalised implementation due

References

Understanding LSTM and its diagrams. *Medium Article*. URL https://medium.com/mlreview/understanding-lstm-and-its-diagrams-37e2f46f1714.

Anjali Thomas Bohlken and Ernest John Sergenti. Economic growth and ethnic violence: An empirical investigation of Hindu-Muslim riots in India. *Journal of Peace Research*, 47(5 Sep 2010):589–600, 2010. ISSN 00223433. doi: 10.1177/0022343310373032.

Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. pages 1–9, 2014. URL http://arxiv.org/abs/1412.3555.

Paul Collier and Anke Hoeffler. On incidence of civil war in Africa. *Center for Development Research (ZEF Bonn)*, 1 (December):1–27, 2000. ISSN 1098-6596. doi: 10.1017/CBO9781107415324.004.

Mihai Catalin Croicu and Ralph Sundberg. UCDP Georeferenced Event Dataset Codebook Quick Start Guide. (October), 2016.

Zhiyong Cui, Ruimin Ke, and Yinhai Wang. Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction. pages 22–25, 1 2018. URL http://arxiv.org/abs/1801.02143.

James D. Fearon and David D. Laitin. Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(1):75–90, 2003. ISSN 15375943. doi: 10.1017/S0003055403000534.

Jonas Gehring, Michael Auli, David Grangier, Denis Yarats, and Yann N. Dauphin. Convolutional Sequence to Sequence Learning. 2017. URL http://arxiv.org/abs/1705.03122.

Alex Graves. Supervised Sequence Labelling. In *Supervised Sequence Labelling with Recurrent Neural Networks*, chapter 3, 4, pages 18–38. Springer-Verlag Berlin Heidelberg, 2012. doi: 10.1007/978-3-642-24797-2{_}2. URL http://link.springer.com/10.1007/978-3-642-24797-2_2.

Alex Graves. Generating Sequences With Recurrent Neural Networks. pages 1–43, 2013. URL http://arxiv.org/abs/1308.0850.

Weisi Guo, Xueke Lu, Guillem Mosquera Donate, and Samuel Johnson. The Spatial Ecology of War and Peace. *Arxiv Preprint: https://arxiv.org/abs/1604.01693*, 2017(June):1–9, 2016. URL http://arxiv.org/abs/1604.01693.

Hvard Hegre and Nicholas Sambanis. Sensitivity analysis of empirical results on civil war onset. *Journal of Conflict Resolution*, 50(4):508–535, 2006. ISSN 00220027. doi: 10.1177/0022002706289303.

Håvard Hegre, Nils W. Metternich, Håvard Mokleiv Nygård, and Julian Wucherpfennig. Introduction: Forecasting in peace research. *Journal of Peace Research*, 54(2):113–124, 2017. ISSN 14603578. doi: 10.1177/0022343317691330.

Håvard Hegre, Marie Allansson, Matthias Basedau, Michael Colaresi, Mihai Croicu, Hanne Fjelde, Frederick Hoyles, Lisa Hultman, Stina Högbladh, Remco Jansen, Naima Mouhleb, Sayyed Auwn Muhammad, Desirée Nilsson, Håvard Mokleiv Nygård, Gudlaug Olafsdottir, Kristina Petrova, David Randahl, Espen Geelmuyden Rød, Gerald Schneider, Nina von Uexkull, and Jonas Vestby. ViEWS: A political violence early-warning system. *Journal of Peace Research*, 56(2):155–174, 2019a. ISSN 14603578. doi: 10.1177/0022343319823860.

Håvard Hegre, Marie Allansson, Matthias Basedau, Michael Colaresi, Mihai Croicu, Hanne Fjelde, Frederick Hoyles, Lisa Hultman, Stina Högbladh, Remco Jansen, Naima Mouhleb, Sayyed Auwn Muhammad, Desirée Nilsson, Håvard Mokleiv Nygård, Gudlaug Olafsdottir, Kristina Petrova, David Randahl, Espen Geelmuyden Rød, Gerald Schneider, Nina von Uexkull, and Jonas Vestby. ViEWS: A political violence early-warning system. *Journal of Peace Research*, 56(2):155–174, 2019b. ISSN 14603578. doi: 10.1177/0022343319823860.

Sepp Hochreiter. The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(2):107–116, 2003. ISSN 0218-4885. doi: 10.1142/s0218488598000094.

Seungkyun Hong, Seongchan Kim, Minsu Joh, and Sa-kwang Song. PSIque: Next Sequence Prediction of Satellite Images using a Convolutional Sequence-to-Sequence Network. *Arxiv Preprint arXiv:1711.10644*, (Dlps):1–5, 2017. URL http://arxiv.org/abs/1711.10644.

Nils W. Metternich, Shahryar Minhas, and Michael D. Ward. Firewall? or Wall on Fire? A Unified Framework of Conflict Contagion and the Role of Ethnic Exclusion. *Journal of Conflict Resolution*, 61(6):1151–1173, 7 2017. ISSN 0022-0027. doi: 10.1177/0022002715603452. URL http://journals.sagepub.com/doi/10.1177/0022002715603452.

Anirban Mitra and Debraj Ray. Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India. *Journal of Political Economy*, 122(4):719–765, 8 2014. ISSN 0022-3808. doi: 10.1086/676316. URL https://www.journals.uchicago.edu/doi/10.1086/676316.

David Muchlinski, David Siroky, Jingrui He, and Matthew Kocher. Comparing Random Forest with Logistic Regression for Predicting Class-Imbalanced Civil War Onset Data. *Political Analysis*, 24(1):87–103, 1 2016. ISSN 1047-1987. doi: 10.1093/pan/mpv024. URL https://www.cambridge.org/core/product/identifier/S1047198700012055/type/journal_article.

Sean P. O'Brien. Crisis early warning and decision support: Contemporary approaches and thoughts on future research. *International Studies Review*, 2010. ISSN 15219488. doi: 10.1111/j.1468-2486.2009.00914.x.

Ola Olsson. Conflict diamonds. *Journal of Development Economics*, 82(2):267–286, 3 2007. ISSN 03043878. doi: 10.1016/j.jdeveco.2005.07.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0304387806000447.

Pau Rodríguez, Miguel A. Bautista, Jordi Gonzàlez, and Sergio Escalera. Beyond one-hot encoding: Lower dimensional target embedding. *Image and Vision Computing*, 2018. ISSN 02628856. doi: 10.1016/j.imavis.2018.04.004.

Heather Sarsons. Rainfall and conflict: A cautionary tale. *Journal of Development Economics*, 115(July):62–72, 7 2015. ISSN 03043878. doi: 10.1016/j.jdeveco.2014.12.007. URL https://linkinghub.elsevier.com/retrieve/pii/S030438781400159X.

Neha Sharma, Vibhor Jain, and Anju Mishra. An Analysis of Convolutional Neural Networks for Image Classification. In *Procedia Computer Science*, 2018. doi: 10.1016/j.procs.2018.05.198.

Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun Woo. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In *NIPS'15 Proceedings of the 28th International Conference on Neural Information Processing Systems*, pages 802–810, 2015. URL http://arxiv.org/abs/1506.04214.

Nitish Srivastava, Elman Mansimov, and Ruslan Salakhutdinov. Unsupervised Learning of Video Representations using LSTMs. In *ICML'15 Proceedings of the 32nd International Conference on International Conference on Machine Learning - Vol 37*, pages 843–852, 2015. URL http://arxiv.org/abs/1502.04681.

Ralph Sundberg and Erik Melander. Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research*, 50 (4):523–532, 7 2013. ISSN 0022-3433. doi: 10.1177/0022343313484347. URL http://journals.sagepub.com/doi/10.1177/0022343313484347.

Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to Sequence Learning with Neural Networks. *Arxiv Preprint*, pages 1–9, 9 2014. URL http://arxiv.org/abs/1409.3215.

Philip E. Tetlock. *Expert Political Judgment*. Princeton University Press, Princeton, 12 2018. ISBN 9781400888818. doi: 10.1515/9781400888818. URL http://www.degruyter.com/view/books/9781400888818/9781400888818/9781400888818.xml.

Andreas Forø Tollefsen, Håvard Strand, and Halvard Buhaug. PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, 49(2):363–374, 2012. ISSN 00223433. doi: 10.1177/0022343311431287.

Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 07-12-June:3156–3164, 2015. ISSN 10636919. doi: 10.1109/CVPR.2015.7298935.

Nils B. Weidmann. Violence "from above" or "from below"? The role of ethnicity in Bosnia's civil war. *Journal of Politics*, 2011. ISSN 00223816. doi: 10.1017/S0022381611000831.

Nils B. Weidmann and Michael D. Ward. Predicting conflict in space and time. *Journal of Conflict Resolution*, 54(6):883–901, 2010. ISSN 15528766. doi: 10.1177/0022002710371669.