

Contextual LSTM 기반 변분 오토인코더를 이용한 이동 경로 예측

조광호*, 차재혁*

한양대학교 컴퓨터소프트웨어학과*

whrhkdghkd@hanyang.ac.kr, chajh@hanyang.ac.kr

Trajectory Prediction by Using Contextual LSTM based Variational AutoEncoder

KwangHo Cho*, JaeHyuk Cha*

Dept. of Computer Software, HanYang University*

요 약

스마트폰, GPS 장비, 위치 기반 소셜네트워크의 발달로 방대한 이동 경로 데이터 수집이 가능하게 됐다. 이를 통해 다양한 분야에서 GPS 데이터를 가지고 사람의 이동성을 분석하고 POI를 예측하는 기회가 많아졌다. 실생활에서 사람의 이동성은 다양한 상황에 영향을 받지만, 실제 GPS 데이터는 위치, 시간 정보의 수준이다. 따라서 다양한 상황을 내재하는 정보가 사람의 이동성 분석과 POI 예측에 필요하다. 본 논문에서는 POI의 순위, 사용자의 POI 활동, 카테고리 선호도 같은 맥락적 특징을 이용하여 이에 관련된 상황에 맞는 POI 시퀀스를 예측하는 Contextual LSTM 기반 딥 러닝 기법을 제안한다. Contextual LSTM은 사람의 이동성에 영향을 주는 시퀀스의 맥락적 특징을 모델에 통합하기 위해 LSTM을 확장한다. 제안된 기법은 HITS 알고리즘과 여러 제약조건 기반으로 추출한 맥락적 특징별로 딥 러닝 모델에 통합하여 각각 POI 시퀀스를 검출했으며, 다양한 맥락적 특징에 대해서 공공 데이터와 수집한 데이터로 평가하였다.

1. 서론

스마트폰, GPS 장비, 위치기반 소셜네트워크(Location Based Social Network, LBSN)의 발달로 인해 수집되는 이동 경로(trajecory) 데이터의 양이 급증하고 있다. 방대한 데이터로 인해 사용자의 이동성을 발견하고 이동 경로를 예측하는 기회가 많아졌다. 사람의 이동성 예측은 목적에 따라 사용자에게 여행 경로 추천, 지능형 교통수단, 도시 계획, 이동성 관리 등 다양한 분야에서 활용된다[1].

최근 몇 년 동안 딥 러닝 분야에서 다양하고 많은 이동 경로 예측 기법이 연구되었지만 대부분 명시적인 과거의 시간, 위치 데이터를 활용하여 사용자의 단일적인 위치 또는 맥락상 의미가 없는 임의의 이동 경로를 예측하는 경우가 많다. 따라서 다양한 상황에 큰 영향을 받는 현실 세계의 이동성과는 많이 다른 결과를 보여준다. 대부분 사용자는 지역적으로 가치가 있는 핫스팟(Hot spot)이나 사용자의 선호에 따라 상황에 맞는 예측된 이동 경로가 필요하다.

본 논문에서는 위치, 시간 정보뿐만 아니라 전체이동 경로에 적용되는 사람의 이동성에 영향을 미치는 새로운 특징 (1) 지역 POI(Point of Interest) 순위점수, (2) 활동 관련 사용자 POI 선호도, 그리고 (3) 카테고리 관련 사용자 POI 선호도를 추출하고, 이와 같은 맥락적 특징들을 통합하여 POI 시퀀스를 예측하는 Contextual LSTM 기반 딥 러닝 모델 기법을 제안한다. 제안된 모델은 사용자의 이동 경로 데이터로부터 더욱 풍부한 정보를 인코딩하고 이를 통해 사용자의 다음 이동 경로를 예측한다. 실제로, 새로운 특징을 통합한 모델은 기존의 모델보다 우수한 예측 성능을 나타낸다. 이후 본 논문의 구성은 다음과 같다.

제2장에서는 Contextual LSTM 기반 딥 러닝 모델에 관한 용어 및 문제 정의와 선행 연구에 대해 다루고, 제3장에서는 새로운 특징들과 모델의 구성도와 각 과정에 대해 설명한다. 제4장에서는 모델에 대한 성능을 비교한다. 마지막 장에서는 결론 및 향후 연구 방향에 대해 기술한다.

2. 용어 및 문제 정의 & 관련 연구

이 장에서는 Contextual LSTM 기반 딥 러닝 모델에 관한 용어 및 문제를 정의하고, 선행 연구에 대해 소개한다.

2.1 용어 및 문제 정의

정의 1. 이동 경로(Trajectory): 각 POI를 l_i 라고 하면, POI 집합들은 $\{l_1, l_2, \dots, l_n\}$ 가 되고, 이를 이동 경로 T_i 라고 한다. 이동 경로 안의 POI들은 시간순의 연속적인 집합이다.

정의 2. 맥락 (Context): 맥락이란 현재 POI에서 다음에 올 POI에 직접적인 또는 간접적인 영향을 미치는 현재 및 이전의 시나리오(senario)를 나타낸다. 예를 들어, 이전 및 현재 장소의 시간/거리 차이, 인기도, 카테고리 등으로 고차원 벡터로 표현될 수 있다.

정의 3. 이동 경로 예측 문제(Trajectory Prediction Problem): $T_i = \{l_1^1, l_2^1, \dots, l_n^1\}$ 이라는 N개의 POI로 이루어진 이동 경로 T_i 가 주어졌을 때, 다음 이동 경로인 $T_{i+1} = \{l_{i+1}^1, l_{i+1}^2, \dots, l_{i+1}^M\}$ 를 예측한다.

2.2 관련 연구

이동 경로 예측에 대한 선행 연구로는 Wang이 제안한 [2]이 있다. [2]에서는 이동 경로 데이터의 전반적인 정보를 인코딩하기 위해 자연어 처리학습(Natural Language Processing, NLP)에 뛰어난 효과를 보여준 Long Term Short Term Memory(LSTM)[3] 순환 신경망과 Seq2Seq를 사용하여 다중 사용자들의 전반적인 이동 경로를 예측하였지만 길이가 긴 이동 경로 입력에는 장기의 존성 문제에 대한 한계가 있었다. 이 문제를 해결하기 위해 어텐션 메커니즘[5]을 적용한 모델[6]도 제안되었다.

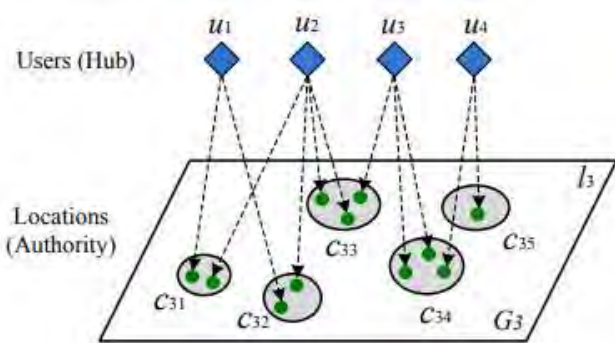
이 외에도 이동 경로 안의 각 데이터끼리의 시간 차이, 거리 차이를 나타내는 위치적 맥락 정보를 변분 오토인코더(Variational Auto Encoder, VAE)[7]를 사용하여 통합하여 인코딩하고 학습한 연구[8][9]가 있다. [8][9]에서 제안한 맥락 학습은 특정 주제에 대해 어떻게 모델링하는 것에 따라 영향력 있는 결과를 나타낼 수 있다고 했다.

3. 제안 기법

3.1 특징점(Features)

이동 경로 데이터로부터 사람의 이동성에 영향을 미치는 3가지 특징을 추출했고, 추출한 특징별로 모델에 통합하여 학습하였다.

3.1.1 지역 POI 순위 점수(Authority of POI): 실제로 위치 예측은 특정 지역의 크기 단위로 행해지는 경우가 대부분이다. 특정 지역이라는 공간에 관해 사람들마다 지식의 정도가 다르며, 또한 POI들의 인기도 또한 다르다. 게다가 그 지역의 경험이 많은 사람이 특정 POI에 가는 것이 더 큰 영향을 미칠 수 있다. 즉, 사람들의 특정 지역에 대한 지식과 그 지역의 POI들의 흥미도는 상호증강 관계를 가진다. 그래서 우리는 그래프 기반 랭킹 알고리즘인 HITS 알고리즘[10]을 사용하여 Authority 점수를 구한다. 단순히 모든 사용자들의 특정 지역의 POI에 대한 방문 횟수와 빈도수로 POI의 인기를 결정하는 게 아니라 특정 지역의 여행 경험이 많은 사람의 방문 횟수와 빈도수에 더욱 영향을 받은 POI들의 순위 점수를 구한다. 그림1은 특정 지역에 대한 POI들(Authorities)과 사용자들(Hubs)의 관계를 나타낸다.



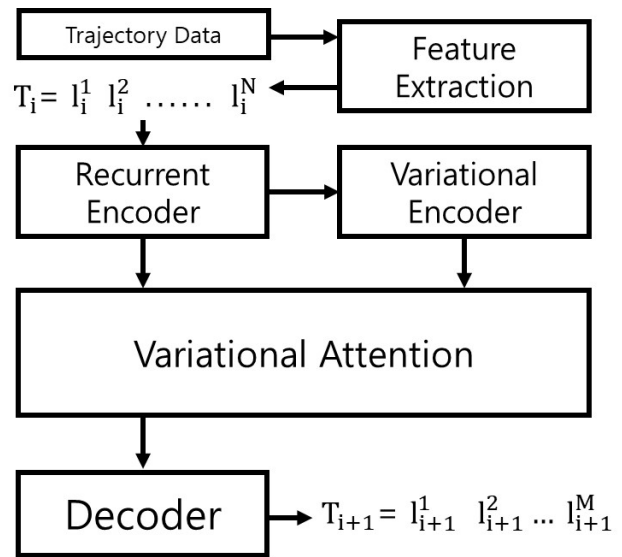
(그림 1) Authority와 Hub의 관계.

3.1.2 개인별 POI 카테고리 선호도 점수(Preference score of POI Category): [11]에서 제안한 특징으로, 각 POI의 평균 체류 시간과 횟수, 해당 POI와 같은 카테고리의 POI들의 평균 체류 시간과 횟수, 그리고 제약조건(거리, 시간)을 이용하여 사용자의 POI에 대한 개인적인 카테고리 선호도 점수를 구한다.

3.1.3 개인별 POI 활동 선호도 점수(Activity score of POI Activity): 개인별 POI 카테고리 선호도 점수와 마찬가지로 사용자가 각 POI에서 한 활동에 대해서 평균 체류 시간과 횟수와 제약조건을 이용하여 사용자의 POI에 대한 활동 선호도 점수를 구한다.

3.2 시스템 설계

아래 그림 2는 현재 순간까지의 이동 경로 데이터로부터 다음 이동 경로를 예측하기 위한 모델의 구성도이다. 최근 자연어 처리 학습에 큰 효과를 보여준 인코더-디코더 구조의 변분 오토인코더(VAE) 모델을 사용했으며, 긴 입력 시퀀스에 대한 장기의존성 문제를 해결하기 위해 어텐션 메커니즘을 적용했다. 각 인코더, 디코더는 LSTM 순환 신경망으로 이루어져 있다.



(그림 2) 제안 모델 구조.

구체적으로 설명하면, 사용자들의 이동 경로 데이터들로부터 위치 정보뿐만 아니라 사람의 이동성에 영향을 미치는 새로운 특징점을 파악하기 위한 특징 추출(Feature Extraction)이 요구된다. 추출된 특징은 지역의 POI 순위 점수, POI 활동 선호도 점수, POI 카테고리 선호도 점수이며, 이 특징들은 모델의 입력이 될 이동 경로 데이터에 매핑(Mapping)된다. 순환 인코더(Recurrent Encoder)는 입력 시퀀스로부터 각 순간의 은닉 상태(hidden state)를 추출하고, 이동 경로 데이터의 맥락적인 정보를 통합하여 전반적인 정보를 나타내는 컨텍스트 벡터(context vector)를 인코딩한다. 다음으로, 변분 인코더(Variational Encoder)는 앞에서 구한 컨텍스트 벡터를 확률 분포인 잠재변수(latent variable)로 변환한다. 동시에, 순환 인코더의 각 순간의 은닉 상태에서 어텐션 벡터(attention vector)를 추출하고 마찬가지로 확률 분포인 잠재변수로 변환한다. 마지막으로, 디코더(Decoder)는 두 개의 잠재변수를 사용하여 다음에 올 가장 확률이 높은 이동 경로를 생성한다.

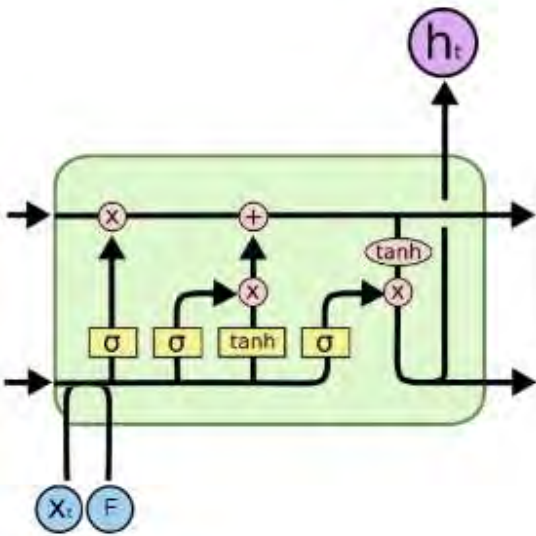
3.3 이동 경로 맥락 인코딩(Trajectory Context Encoding)

3.2.2 Recurrent Trajectory Encoder

순환 인코더는 양방향 LSTM[12] 순환 신경망으로 구성되어 있다. 양방향 은닉층의 결합에 의해 POI의 은닉 상태를 얻을 수 있다. 이동 경로의 맥락적인 정보들을 은닉

상태에 통합하는 것이 사용자의 이동성을 표현하는데 훨씬 도움이 된다. 그러나, 기본적인 LSTM(Vanilla LSTM)은 POI와 관련된 맥락적인 정보를 처리하는데 적합하지 않기 때문에 확장된 LSTM을 제안한다. 그림3은 새롭게 확장된 LSTM 구조이다. LSTM의 메모리 셀 내/외부부의 정보 흐름을 조절하는 기존의 입력/출력/망각 게이트 안에 추출한 특징 벡터를 추가하여 Contextual LSTM으로 확장한다. 식1은 확장된 LSTM 게이트의 동작 식이다.

확장된 LSTM은 순간마다 이전 순간의 POI와 다음 순간의 POI의 정보를 포함하고 있는 은닉 상태를 추출하고, 마지막 순간의 은닉 상태를 이동 경로 데이터의 맥락적인 정보를 통합하여 전반적인 정보를 나타내는 고정된 크기의 하나의 컨텍스트 벡터(context vector) c 로 인코딩한다.



(그림 3) 확장된 LSTM 구조.

$$\begin{aligned}
 i_t &= \sigma(W_i x_t + U_i h_{t-1} + V_i F + b_i) \\
 f_t &= \sigma(W_f x_t + U_f h_{t-1} + V_f F + b_f) \\
 o_t &= \sigma(W_o x_t + U_o h_{t-1} + V_o F + b_o) \\
 c_t &= f_t \circ c_{t-1} + i_t \circ \sigma(W_c x_t + U_c h_{t-1} + V_c F + b_c) \\
 h_t &= o_t \circ \tanh(c_t)
 \end{aligned}$$

(식 1) 확장된 LSTM의 게이트 동작 식.

3.2.3 Variational Trajectory Encoder

변분 인코더를 앞에서 구한 고정된 컨텍스트 벡터를 확률 분포인 잠재변수 z 로 인코딩하는 데 사용한다. 생성된 잠재변수는 랜덤 샘플(Random Sample)을 사용하여 디코더에서 다양한 데이터를 재구성할 수 있다[7].

3.2.4 Variational Attention

일반적인 어텐션 메커니즘은 입력 시퀀스와 정답 시퀀스에 대한 정렬 관계를 자동으로 학습한다. 구체적으로, 입력 시퀀스의 모든 순간마다 해당 순간의 정답과 어느 정도 관련이 있는지 확률 값을 계산하고, 계산된 확률 값을 각 입력 시퀀스의 은닉 상태에 가중치 합해주어 어텐션 벡터를 구한다[5]. 어텐션 벡터 또한 컨텍스트 잠재변수와 마찬가지로 확률 분포인 잠재변수로 인코딩한 후 디코더로 전달된다.

3.2.5 Trajectory Decoders

위에서 언급하였듯이, 디코더는 현재 이동 경로가 주어졌을 때, 어텐션 잠재변수 c 와 컨텍스트 잠재변수 z 를 이용하여 다음에 올 이동 경로를 예측한다. 두 잠재변수와 이전 순간 디코더 은닉 상태를 사용하여 다음에 오는 POI를 예측하고 식2의 조건부 확률 식이 최대화되도록 학습한다.

$$p(\{l_{i+1}^1, l_{i+1}^2, \dots, l_{i+1}^M\}) = \prod_{t=1}^M p(l_{i+1}^t | l_{i+1}^{<t}, c, z)$$

(식 2) 디코더의 조건부 확률 식.

4. 평가

이 섹션에서는 이동 경로 예측 문제에 대한 데이터셋, BaseLine, 평가 메트릭, 실험결과에 대해 설명한다.

4.1 데이터 셋

본 논문에서 구하려는 이동 경로의 새로운 맥락적 특징 추출과 관련 있는 속성을 가지고 있고 잘 정리된 특정 지역의 이동 경로 데이터를 사용하여 평가했다: Weeplaces 공공 데이터, 서울시 성동구 데이터

Weeplaces[13] 데이터는 위치 기반 소셜네트워크 Weeplaces에서 제공한 오픈 데이터로서, 15,799명의 사용자가 등록한 위치 정보가 수집되었으며, 7,658,368개의 체크인 위치가 포함되어 있다. 각 장소의 카테고리가 10개로 분류되었으며, 장소에서 한 활동을 96개로 분류했다.

서울시 성동구 데이터는 위치수집 앱인 트래카(Traccar)[14]를 사용해서 2주 기간의 성동구 주변 사람들의 위치를 수집한 데이터이다. 각 장소의 카테고리를 9개로 분류했다.

모든 데이터 셋에 대해서, 훈련을 위해 90%를 무작위 선택하여 훈련데이터로 사용하고, 나머지 10%는 테스트 데이터로 사용하였다.

4.2 BaseLines

본 논문에서는 예측한 이동 경로의 정확성과 방문순서의 타당성을 측정하기 위해 두 가지 메트릭(Metrics)을 사용하여 평가하고, 각 특징별로 통합한 제안된 모델과 기존의 이동 경로 예측 모델들과 비교 한다.

VAE(Variational AutoEncoder): 인코더-디코더로부터 생성되는 데이터의 확률 분포를 잠재변수로 생성하고 학습하는 모델

VAE with Attention: VAE 모델에 어텐션 메커니즘을 적용한 모델

CATHI[9]: VAE를 사용하여 POI의 시간/거리 차이만 맥락적 특징으로 통합하여 학습하는 예측 모델.

위의 2개의 모델은 기존의 LSTM을 그대로 사용한 모델들이고, **CATHI**는 LSTM을 확장하여 맥락적 특징을 통합한 모델이다.

4.3 Metrics.

우리는 F1 점수(F1 score)와 pairs-F1 점수(pairs-F1 score)[15]를 사용하여 이동 경로 예측을 평가했습니다. F1 점수는 예측된 이동 경로의 POI들이 실제 이동 경로의 POI들과 적합한지에 대한 점수를 나타내고, pairs-F1 점수는 이동 경로의 POI 순서가 적합한지를 확인한다. pairs-F1 점수는 순서대로 POI들의 쌍의 F1 점수를 사용해서 계산한다. 다음 식3은 pairs-F1을 구하는 식으로, P_{pairs} 와 R_{pairs} 는 순서대로 POI 쌍들의 정밀도(Precision)

와 재현율(Recall)을 나타낸다.

$$pairs - F_1 = \frac{2P_{pair}R_{pair}}{P_{pair} + R_{pair}}$$

(식 3) pairs-F1 식

4.4 실험 결과

전반적인 성능

< 표1 > Baseline과 제안 모델 실험 결과

		Trajectory Prediction	
Methods		F1	Pairs-F1
We eplace	VAE	0.599±0.174	0.347±0.209
	VAE+Att	0.624±0.159	0.365±0.166
	CATHI	0.749±0.123	0.473±0.150
	Our Model (Feature1)	0.790±0.108	0.485±0.138
	Our Model (Feature2)	0.862±0.087	0.615±0.105
	Our Model (Feature3)	0.829±0.098	0.560±0.117
성동구 위 치 데이터	VAE	0.409±0.143	0.223±0.342
	VAE+Att	0.420±0.122	0.248±0.304
	CATHI	0.557±0.102	0.358±0.209
	Our Model (Feature1)	0.621±0.094	0.365±0.166
	Our Model (Feature2)	0.653±0.087	0.473±0.130

다음 표 1은 각 데이터에 대한 예측 결과의 F1 점수와 pairs-F1 점수입니다. 논문에서 제안된 모델들은 확장한 Contextual LSTM 기반 VAE모델에 각각 특징1(POI 순위), 특징2(POI 선호도(카테고리)), 특징3(POI 선호도(활동))을 통합한 것이다. F1, pairs-F1 점수 측면에서 전반적으로 모든 Baseline보다 맥락적 특징을 통합한 모델이 그렇지 않은 모델에 비해 성능이 크게 개선된 것을 볼 수 있다. 기존 VAE보다 어텐션 메커니즘을 적용한 VAE가 좀 더 우수한 성능을 보이고, 시간/거리 차이만 맥락적 특징으로 통합하여 학습시킨 CATHI는 더욱 향상이 되었다. 즉, 제안된 모델의 맥락적인 정보 인코딩이 이동 경로의 정보를 잘 내재하며, 예측에 큰 영향을 미친다는 뜻이며, 우리가 추가로 넣은 새로운 맥락적 정보 중 사용자 POI 선호도(카테고리) 특징2가 사용자의 이동성에 가장 큰 영향을 미친다는 것을 알 수 있다. 이유는 특징1(Authority 점수)은 사람들의 특정 지역 POI의 방문 수를 기반으로 구한 것으로, 개인화 측면보다는 전반적인 관계를 나타내고 특징2(사용자 POI 선호도(카테고리)), 특징3(사용자 POI선호도(활동))은 개인에 관한 POI의 선호도를 구한 것으로 이동 경로를 예측하는 각 사용자 별 이동성을 효율적으로 모델링한다.

이 평가 결과는 어텐션 메커니즘이 인코더-디코더 구조 모델을 향상시키는데 기여를 하며, LSTM의 확장에 대한 주장을 뒷받침하며 맥락적 특징을 통합한 LSTM의 변형이 사람의 이동성 모델링에 효율적이고, 개인별 이동성을 내재하는 맥락적 특징들이 이동 경로 예측에 큰 영향을 미친다는 것을 알 수 있다.

또한, 제안된 모델이 다른 방법에 비해 표준편차의 값이 낮은 것을 보았을 때, 더 안정적이라는 걸 알 수 있다.

5. 결론 및 미래 연구

본 논문에서는 맥락적 정보에 의존하고, 추출한 새로운 특징을 확장된 Contextual LSTM 기반 딥러닝 모델에 통합했다. 두 개의 데이터 셋으로 제안된 모델을 평가하고 맥락적인 정보를 통합하는 것이 사람의 이동성을 표현하는 데 도움이 될을 보여준다. 제안된 모델은 더 큰 데이터 셋에서 약간 더 나은 성능을 보였다. 그리고 HITS 알고리즘, POI 선호도 식을 사용해서 얻은 새로운 특징들은 이동 경로 데이터를 나타내는 데 좋은 지침이라는 것을 보여주며, 새롭게 확장한 LSTM으로 더 유의미한 정보를 내재한 컨텍스트 벡터를 제공한다는 걸 보여준다. 이동 경로 데이터로부터 추출할 수 있는 흥미로운 맥락적 특징은 많이 있다. 앞으로 사용자의 선호도 및 장소의 인기도를 정의하기 위해 텍스트 정보(예: 태그, 팁 및 리뷰)와 시각적 정보(예: 장소 사진) 또한 통합해보려고 한다.

사사

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2018R1A5A7059549).

참고문헌

- [1] Z. Kai, S. Tarkoma, S. Liu, and H. Vo, "Urban human mobility data mining: An overview," in Proc. IEEE Int. Conf. Big Data, Dec. 2017, pp. 1911 - 1920.
- [2] W. Chujie, M. Lin, and H. Zhang, "Exploring Trajectory Prediction Through Machine Learning Methods" in IEEE Access, July. 2019, pp.101441 - 101452.
- [3] Sepp Hochreiter, Jürgen Schmidhuber, "Long short-term memory". Neural Computation. 9 (8), 1997, 1735 - 1780.
- [4] Kyunghyun Cho, Bart van Merriënboer, Dzmitry Bahdanau and Yoshua Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation" in EMNLP, Jun 2014.
- [5] Dzmitry Bahdanau, Kyunghyun Cho and Yoshua Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate" in ICLR, 2015.
- [6] Jie Feng, Yong Li, Chao Zhang, Funing Sun, Fanchao Meng, Ang Guo, and Depeng Jin, "DeepMove: Predicting Human Mobility with Attentional Recurrent Networks". In WWW, 2018.
- [7] Diederik P. Kingma and Max Welling, "Auto-Encoding Variational Bayes", in cs.LG, May 2014.
- [8] Jarana Manotumruksa, Craig Macdonald, and Iadh Ounis, "A Contextual Attention Recurrent Architecture for Context-Aware Venue Recommendation", In SIGIR. 2018.
- [9] Fan Zhou, Xiaoli Yue and Goce trajceviski, "Context-aware Variational Trajectory Encoding and Human Mobility Inference", in WWW, May 2019.
- [10] Christopher D. Manning, Prabhakar Raghavan & Hinrich Schütze, "Introduction to Information Retrieval". Cambridge University Press. Retrieved, 2008.
- [11] Ramesh Baral, Tao Li, XiaoLong Zhu, "CAPS: Context Aware Personalized POI Sequence Recommender System", in Information Retrieval (cs.IR), Mar 2018.
- [12] Schuster, Mike, and Kuldip K. Paliwal. "Bidirectional recurrent neural networks." Signal Processing, IEEE Transactions, 1997, 2673-2681.
- [13] X. Liu, Y. Liu, K. Aberer, C. Miao, "Personalized point-of-interest recommendation by mining users' preference transition", in CIKM, 2013, pp.733 - 738.
- [14] <https://www.traccar.org/documentation/Traccar>, Traccar Documentation.
- [15] Dawei Chen, Cheng Soon Ong, Lexing Xie, "Learning Points and Routes to Recommend Trajectories", in CIKM '16, 2016.