

uncertainty, 즉 불확실성을 측정하는 것은 답러닝에 있어서 중요한 과제가 되었다. 이런 분야를 Uncertainty Quantification(UQ)로 따로 분류도 해놓은 것 같다. 기존에 이를 위한 시도가 있었는데, 그중 BNN은 모수 추정이 intractable하다는 점과 모수 공간 크기가 너무 커서 단점이 있었다. 또 다른 방법으로 앙상블을 통해서 구하는 방식이 있는데, 각기 다른 모델을 여러번 학습한다는 점이 비효율적이다. 그래서 이 논문은 두가지 종류의 불확실성을 같이 학습할 수 있고 분리해서 볼수 있는, 그리고 직관적이고 사전확률분포를 정의하지 않아도 되는 'SDE-Net'을 제시한다. 그리고 classification, regression 모두 사용가능하다.

aleatoric uncertainty: 우리가 측정할 수 없는 원인에 의해 일어나는 오류

예를 들면, 화살 쏘기에서 화살의 궤적..

이 논문에서는 데이터 셋 내에서의 오차.

epistemic uncertainty: 우리가 알고 있지만 실재적으로 무시하고 사용하지 않는 오류

예를 들면, 정격가속도 측정설비에서 공기저항을 무시..

이 논문에서는 과충치 절제를 위해 일어나는 오류

이 단점을 해결하려는 시도를 Variational Inference와 MCMC가 있다.

< Content >

모델을 구현할 때 Neural ODE를 사용한다. 그런데 기존 Neural ODE는 deterministic 하기 때문에 불확실성을 잡아낼 수 없다. 그래서 여기에 Brownian motion을 접목시킨다. 모델은 두가지 part로 나뉘어져 있다.

① drift net 은 predict를 이끄는 main net 이다.

② diffusion net은 위의 예측값에 노이즈를 더하는 것이다.

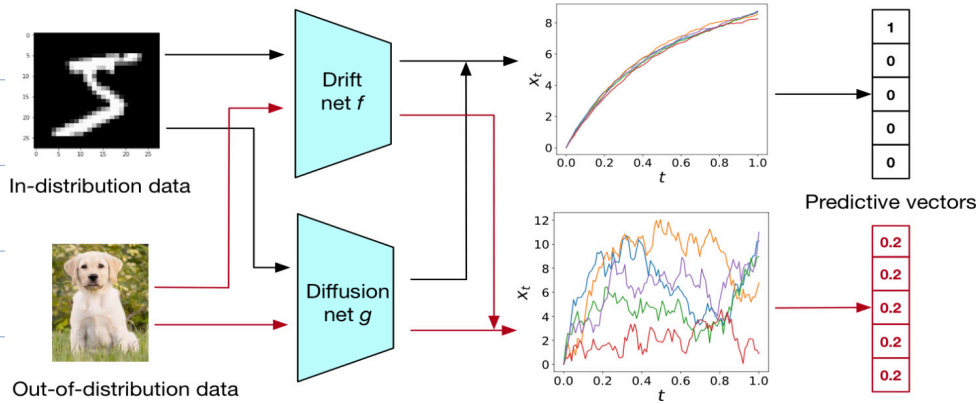
요약하면 drift net이 정확도를 높이는 역할을 하는 만큼 diffusion net이 모델의 불확실성을 더해준다.

Neural ODE : $dx_t = f(x_t, t) dt$

SDE-Net : $dx_t = \underbrace{f(x_t, t)}_{\text{drift net}} dt + \underbrace{g(x_t, t)}_{\text{diffusion net}} dW_t$
 (drift net, diffusion net, 가우시안 노이즈)

(W_t 는 Brownian motion)

만약 train data가 충분하다면 g 는 작을 것이다. 반대를 부족하다면 g 값은 클 것이다 (= 불확실해진다)



위 그림에서 보이는 것처럼 정상적인 data가 들어가면 불확실이 작아서 f 를 따라서 가는 것을 볼 수 있다. 하지만 비정상적 data가 들어가면 불확실이 커져서 f 를 따라가더라도 이상한 결과가 나오는 것을 알 수 있다.

objective function은 다음과 같다.

drift net이
 잘 작동할 수 있도록

$$\begin{aligned} & \min_{\theta_f} E_{x_0 \sim P_{\text{train}}} E(L(x_T)) + \min_{\theta_g} E_{x_0 \sim P_{\text{train}}} g(x_0; \theta_g) \quad \text{정상 데이터는 불확실이 작게끔} \\ & + \max_{\theta_g} E_{\tilde{x}_0 \sim P_{\text{OOD}}} g(\tilde{x}_0; \theta_g) \quad \text{비정상적 데이터는 불확실이 크게끔} \\ \text{s.t. } & dx_t = \underbrace{f(x_t, t; \theta_f)}_{\text{drift neural net}} dt + \underbrace{g(x_0; \theta_g)}_{\text{diffusion neural net}} dW_t, \end{aligned}$$