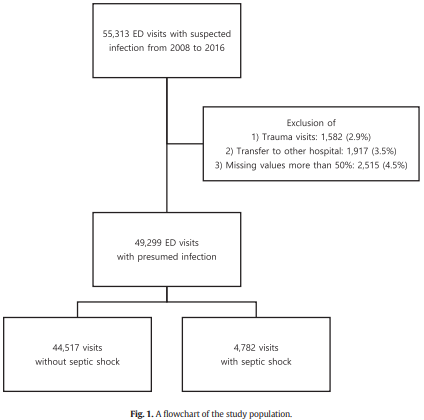
Machine learning for prediction of septic shock at initial triage in ED

Joonghee Kim et al.

Obj: ML이 conventional scoring system보다 성능이 뛰어나다.

Data: A single-center observational study.

1. Aged >= 20years in ED
2. A suspected infection
   1. Taking cultures
   2. A prescription of systemic antibiotics within 24h of ED arrival
3. Septic shock definition (SEPSIS-3) with **adequate volume resuscitation** within 3 hours of ED arrival(assumption)  
   *Adult patients with septic shock can be identified using the clinical criteria of hypotension* ***requiring use of vasopressors to maintain mean blood pressure of 65 mm Hg or greater*** *and having* ***a serum lactate level greater than 2 mmol/L*** *persisting* ***after adequate fluid resuscitation****.* (Shankar-Hari M, Phillips GS, Levy ML, Seymour CW, Liu VX, Deutschman CS, et al. Developing a new definition and assessing new clinical criteria for septic shock: for the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA 2016;315:775–87. <https://doi.org/10.1001/jama.2016.0289>.)
   1. CVI >= 2 (3h <= ED treatment <= 24h) with lactate >= 2
   2. If no lactate, CVI >= 2 (12h <= ED treatment <= 24h)

Singer M, Deutschman CS, Seymour C, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (Sepsis3). JAMA 2016;315:801–10. <https://doi.org/10.1001/jama.2016.0287>.

* + - hypotension should be denoted as a mean arterial pressure less than 65 mm Hg
    - an elevated lactate level is reflective of cellular dysfunction in sepsis (>2 mmol/L [18 mg/dL])
    - Criteria for “adequate fluid resuscitation” or “need for vasopressor therapy” could not be explicitly specified because these are highly user dependent, relying on variable monitoring modalities and hemodynamic targets for treatment.
    - The first database interrogated was **the Surviving Sepsis Campaign’s international multicenter** registry of 28 150 infected patients with at least 2 SIRS criteria and at least 1 organ dysfunction criterion.

내 생각

본 논문 자체가 시사하는 바는 많지는 않다고 생각한다.  
다만 다음과 같은 것들은 유용할 듯하다.

텍스트이(가) 표시된 사진

자동 생성된 설명

**SBP, DBP, AVPU(Glasgow coma score)는 significant한 variable이다**.  
**MLP(Multi-layer Perceptron)보다는 GBM 또는 Ensemble model이 더 성능이 좋다**.

테이블이(가) 표시된 사진

자동 생성된 설명

주목할 것은 GBM, ensemble model.  
다만, AUROC는 수치가 높은 편인데, Sensitivity가 그에 비해 높은 수치는 아니라는 점이 걸린다. (준수하다고는 볼 수 있을 듯) Specificity가 더 높은 것인가? 그건 왜 표기하지 않았을까…

Positive Predicted Value가 너무 낮다. 저 비율은 오히려 FP가 더 많다는 뜻 아닌가?

🡺 다만 이러한 경우, 이는 곧 저자가 Limitation으로 언급한, *this is a single center study, which could represent a concern for generalizability* 라는 부분을 감안하면 어느정도 이해는 가능하다.  
일반적으로 Septic shock의 실험군보다는 대조군이 많을 수 밖에 없기 때문에, 모델도 실험군보다는 대조군을 잘 학습할 수 밖에 없다. 그러므로 양성을 찾는 능력보다는 음성을 찾는 능력 (NPV, Specificity)가 수치가 더 잘 나온다고 볼 수 있다.

만약 그렇다면 여기서 또 한 가지 드는 의문은, 왜 Undersampling 등으로 imbalance problem을 해결하지 않았을까 하는 것이다. 그것만으로 좀 더 좋은 결과가 나올 수 있었을텐데 아쉽게 생각된다.

추가적인 한계점.

1. 우리가 사용하는 데이터는 ICU인 반면, 이것은 ED이다.  
   그러므로 환자들의 중증도가 전반적으로 낮아서 대조군 환자의 상태가 Septic shock 직전인 환자와 극명할 것이다. 이는 오히려 한정적인 결과를 가져올 수 있다. (vital이 아슬하거나, 애매한 환자들)
2. 24시간이내로 한정지은 부분. 만약 하루 이내까지 아무 문제없다가, 이후부터 문제가 발생한 데이터는 모두 버린 것이 되어버린다.
3. SIRS를 사용하지는 않는 부분. 대신 본 논문에서는 이와 관련된 연구 논문들을 몇 편 소개하고 있다.
4. 논문에서 언급하는 한계점
   1. 임상등록부 대신 EHR DB를 이용한 점.
   2. 환자의 기본상태는 예측 변수로서 불필요하였다.
   3. 너무 많은 알고리즘을 테스트한 점. (동의)
   4. Data split이 60:40인 점. Training set의 비율을 더 늘리지 못한 점.