

Proje Raporu: MIMIC-III Veri Seti ile Sepsis Tespiti İçin Keşifsel Veri Analizi

Hazırlayan: [Efe Yarkın Birsin] Tarih: [30.06.2025] Yarışma/Proje Adı: [SepsisGuard]

İçindekiler

- [Giriş: Proje Amacı ve Sepsisin Önemi](#)
- [Veri Setine Genel Bakış ve Analiz Stratejisi](#)
- [Tablo Bazında Detaylı Analiz](#)
 - [3.1. ICUSTAYS.csv: Yoğun Bakım Kalış Bilgileri](#)
 - [3.2. D_ITEMS.csv & D_LABITEMS.csv: Veri Sözlükleri](#)
 - [3.3. INPUTEVENTS_CV.csv & INPUTEVENTS_MV.csv: Tedavi Girdileri](#)
 - [3.4. Diğer Önemli Tablolar \(CHARTEVENTS vb.\)](#)
- [Sonuç ve Sonraki Adımlar](#)

1. Giriş: Proje Amacı ve Sepsisin Önemi

Sepsis, vücudun bir enfeksiyona verdiği düzensiz ve hayatı tehdit eden yanıt sonucu ortaya çıkan, yüksek mortalite oranlarına sahip kritik bir sendromdur. Erken teşhis ve hızlı müdahale, hastanın hayatta kalma şansını önemli ölçüde artırmaktadır.

Bu projenin temel amacı, MIMIC-III yoğun bakım verilerini kullanarak, hastaların sepsis geliştirme riskini YBÜ'ye yatışlarının erken aşamalarında tahmin edebilen bir makine öğrenmesi modeli geliştirmektir. Bu rapor, model geliştirme sürecinin ilk ve en kritik adımı olan **Keşifsel Veri Analizi (EDA)** bulgularını sunmaktadır.

2. Veri Setine Genel Bakış ve Analiz Stratejisi

Bu çalışmada, 40,000'den fazla hastanın verisini içeren MIMIC-III veritabanının, 100 hastayı kapsayan demo versiyonu kullanılmıştır. Demo set, orijinal veri setinin şemasını yansıtmakta ve prototip geliştirme için ideal bir ortam sunmaktadır.

Analiz Stratejimiz: Sepsis tanısı için kullanılan SOFA skoru gibi klinik standartları referans alarak, organ fonksiyon bozukluklarını ve enfeksiyon belirtilerini yansıtan kritik verileri (yaşamsal bulgular, laboratuvar sonuçları, ilaç ve sıvı tedavileri) tespit etmek ve bu verilerin zaman içindeki değişimini modellemektir.

3. Tablo Bazında Detaylı Analiz

Bu bölümde, projemiz için kritik öneme sahip olan her bir tablo tek tek incelenmiş, içerdiği bilgiler, veri kalitesi ve modellemeye olan katkısı kod örnekleri ve görselleştirmelerle ortaya konmuştur.

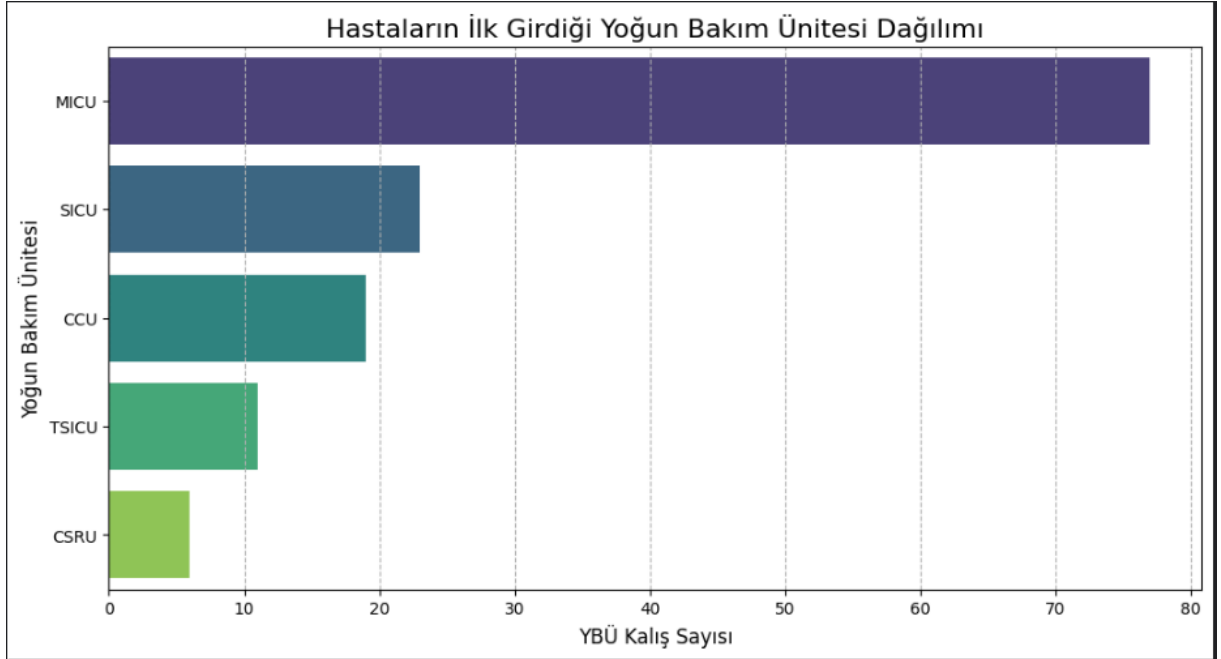
3.1. ICUSTAYS.csv: Yoğun Bakım Kalış Bilgileri

Bu tablo, analizimizin temel çerçevesini oluşturur. Her bir YBÜ kalışını, giriş ve çıkış zamanlarıyla birlikte tanımlayarak zaman serisi analizlerimizin sınırlarını belirler.

	row_id	subject_id	hadm_id	icustay_id	dbsource	first_careunit	last_careunit	first_wardid	last_wardid	intime	outtime	los
0	12742	10006	142345	206504	carevue	MICU	MICU	52	52	2164-10-23 21:10:15	2164-10-25 12:21:07	1.6325
1	12747	10011	105331	232110	carevue	MICU	MICU	15	15	2126-08-14 22:34:00	2126-08-28 18:59:00	13.8507
2	12749	10013	165520	264446	carevue	MICU	MICU	15	15	2125-10-04 23:38:00	2125-10-07 15:13:52	2.6499
3	12754	10017	199207	204881	carevue	CCU	CCU	7	7	2149-05-29 18:52:29	2149-05-31 22:19:17	2.1436
4	12755	10019	177759	228977	carevue	MICU	MICU	15	15	2163-05-14 20:43:56	2163-05-16 03:47:04	1.2938

Column Non-Null Count Dtype

0 row_id 136 non-null int64
1 subject_id 136 non-null int64
2 hadm_id 136 non-null int64
3 icustay_id 136 non-null int64
4 dbsource 136 non-null object
5 first_careunit 136 non-null object
6 last_careunit 136 non-null object
7 first_wardid 136 non-null int64
8 last_wardid 136 non-null int64
9 intime 136 non-null datetime64[ns] 10 outtime 136 non-null datetime64[ns] 11 los 136 non-null float64
dtypes: datetime64[ns], float64(1), int64(6), object(3) memory usage: 12.9+ KB



--- YBÜ Tiplerine Göre Kalış Sayıları ---

first_careunit

MICU 77

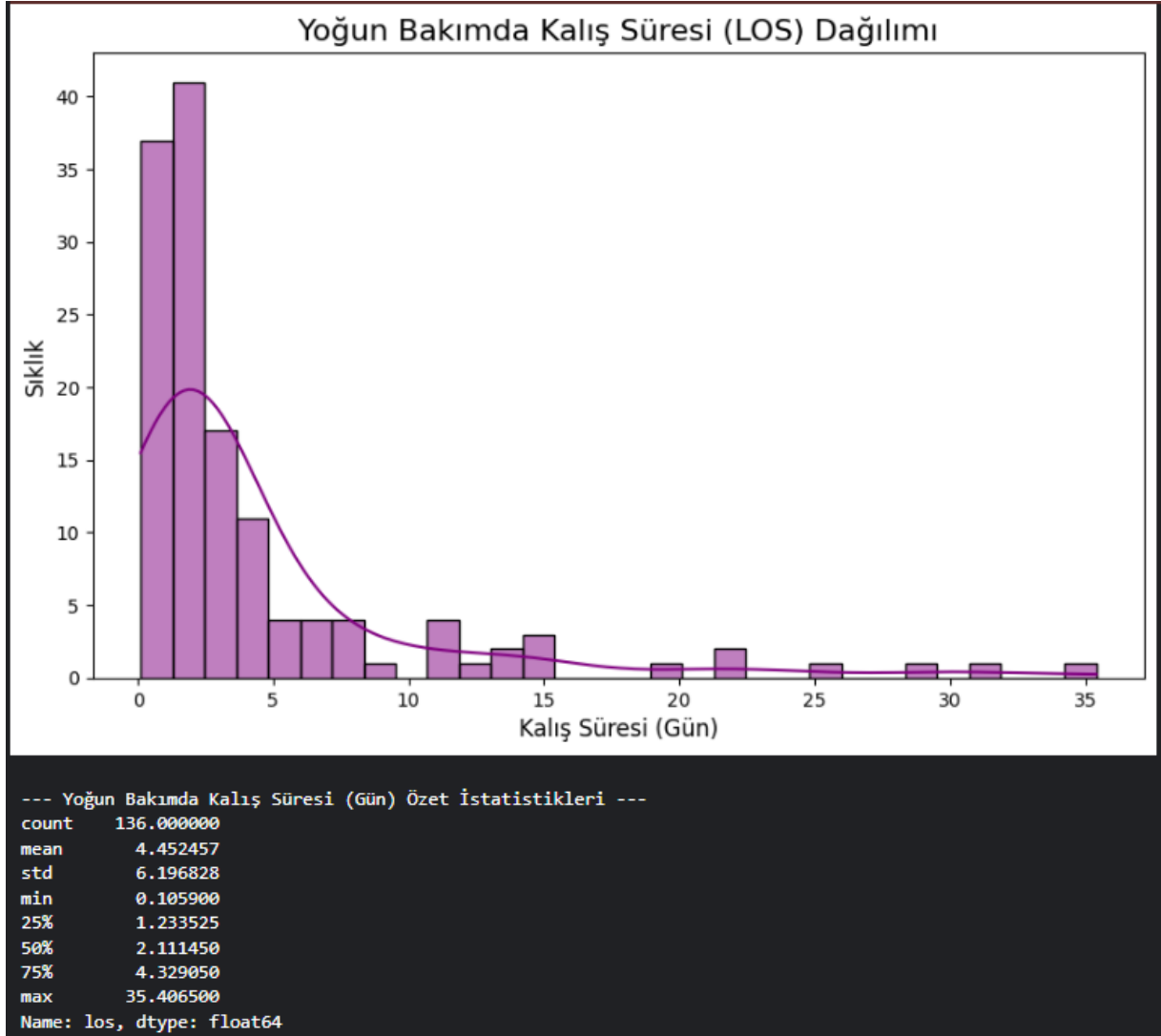
SICU 23

CCU 19

TSICU 11

CSRU 6

Name: count, dtype: int64



🧠 ICUSTAYS.csv: Kullanılabilirlik Özeti ve Proje İçin Rolü

📌 Genel Amaç Bu tablo, analizimizin temelini oluşturan zaman pencerelerini ve analiz birimlerini tanımlar. Hangi verinin, hangi zaman aralığında, hangi hastaya ait olduğunu belirlememizi sağlar.

✅ Model Hazırlığındaki Rolü

Analiz Birimi: Her bir satır (icustay_id), modelimiz için işlenecek bağımsız bir veri örneğidir.

Zaman Çerçevesi: Her bir icustay_id için intime ve outtime sütunları, diğer olay tablolarından (CHARTEVENTS, LABEVENTS vb.) veri çekerken kullanılacak kesin zaman sınırlarını belirler.

Özellik Kaynağı: los (kalış süresi) ve first_careunit (YBÜ tipi) gibi sütunlar, doğrudan model için öznitelik (feature) olarak kullanılabilir.

⚠ Veri Kalitesi Notu Bu tablo genellikle oldukça temiz ve eksiksizdir. Yine de, los değeri negatif olan veya intime >= outtime olan anormal kayıtların kontrol edilip gerekirse analizden çıkarılması iyi bir pratiktir.

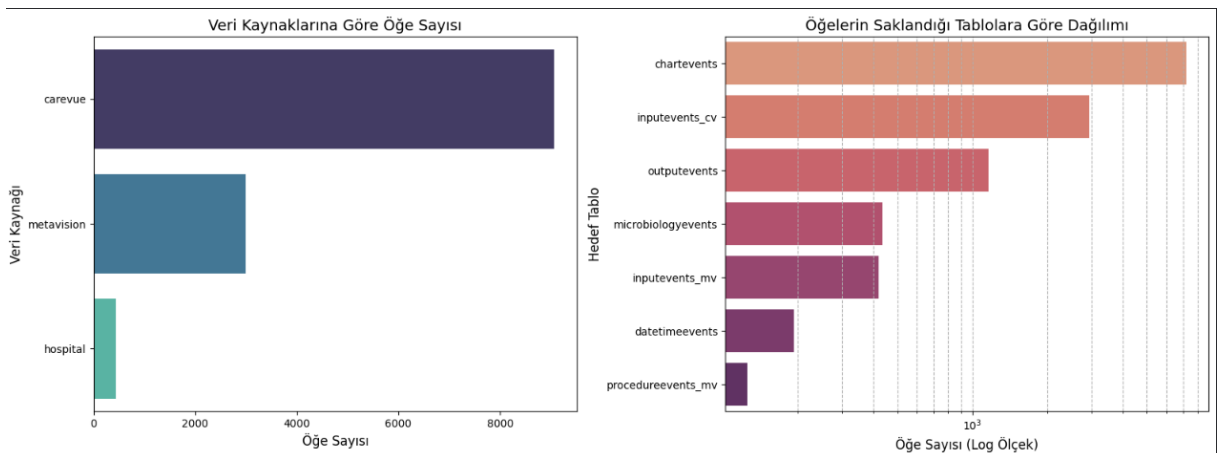
🚩 Sonuç ICUSTAYS.csv, projenin organizasyonel bel kemiğidir. Bu tablo olmadan, farklı tablolardaki milyonlarca veriyi anlamlı, hasta bazlı zaman serilerine dönüştürmek mümkün olmazdı. Analizlerimizin her biri, bu tablodaki icustay_id'ler etrafında şekillenecektir.

3.2. D_ITEMS.csv & D_LABITEMS.csv: Veri Sözlükleri

Bu iki tablo, veritabanının "Rosetta Taşı"dır. Olay tablolarındaki itemid kodlarını anlamlı klinik etiketlere çevirirler. Bu tablolar olmadan diğer olay tabloları kullanılamaz.

ID	Sütun Adı	Anlamı	Proje Bağlantısı
row_id	Satır numarası	Teknik amaçlı, modelde kullanılmaz.	
itemid	Öge (Item) ID'si	EN KRİTİK SÜTUN. Olay tablolarındaki itemid ile eşleştirilerek o verinin ne olduğunu belirler.	
label	Öge Etiket (Açıklama)	itemid 'nin insan tarafından okunabilir karşılığı (örn. "Heart Rate", "Norepinephrine"). Çok önemli.	
dbsource	Veri kaynağı sistemi	carevue veya metavision . Aynı ölçümün farklı ID'lere sahip olabileceğini gösterir, birleştirme için kritiktir.	
linksto	Bağlantılı olduğu tablo	Bu itemid 'nin verisinin hangi olay tablosunda (chartevents , inputevents_mv vb.) olduğunu belirtir. Çok önemli.	
category	Kategori	Ölçümlerin genel sınıfı (örn. "Vital Signs", "Labs", "Medications"). Veri filtreleme için son derece kullanışlıdır.	
unitname	Ölçüm Birimi	Verinin birimi (örn. "bpm", "mmHg", "mg/hour"). Veri standardizasyonu ve normalizasyonu için kritiktir.	
param_type	Parametre Tipi	Verinin sayısal mı (Numeric) yoksa metin mi (Text) olduğunu belirtir.	

	row_id	itemid	label	abbreviation	dbsource	linksto	category	unitname	param_type
0	1	1435	Sustained Nystamus	NaN	carevue	chartevents	NaN	NaN	NaN
1	2	1436	Tactile Disturbances	NaN	carevue	chartevents	NaN	NaN	NaN
2	3	1437	Tremor	NaN	carevue	chartevents	NaN	NaN	NaN
3	4	1438	Ulnar Pulse [Right]	NaN	carevue	chartevents	NaN	NaN	NaN
4	5	1439	Visual Disturbances	NaN	carevue	chartevents	NaN	NaN	NaN



D_ITEMS.csv : Kullanılabilirlik Özeti ve Proje İçin Rolü

🚩 **Genel Amaç** Bu tablo, olay tablolarındaki (CHARTEVENTS , INPUTEVENTS , vb.) ham, kodlanmış veriyi, klinik olarak yorumlanabilir ve modellenenebilir bilgilere dönüştürmek için kullanılan bir **çeviri sözlüğüdür**. Temizlenmesi gereken bir veri seti değil, diğer veri setlerini temizlemek ve hazırlamak için kullanılan bir **araçtır**.

✅ Model Hazırlığındaki Rolü

- **Özellik Seçimi:** Sepsis için önemli olan tüm fizyolojik ölçüm, ilaç ve sıvıların **itemid** 'lerini bu tablodan belirlemek.
- **Veri Filtreleme:** Milyonlarca satırlık olay tablolarını, sadece seçilen bu kritik **itemid** 'leri içerecek şekilde filtreleyerek veri boyutunu dramatik olarak azaltmak.
- **Veri Birleştirme:** Farklı sistemlerden (carevue , metavision) gelen ve aynı anlama gelen ölçümleri (örn. Kalp Hızı) tek bir standart isim altında birleştirmek.
- **Birim Standardizasyonu:** unitname sütununu kullanarak tüm ölçümleri (örn. sıcaklık) standart bir birime çevirmek.

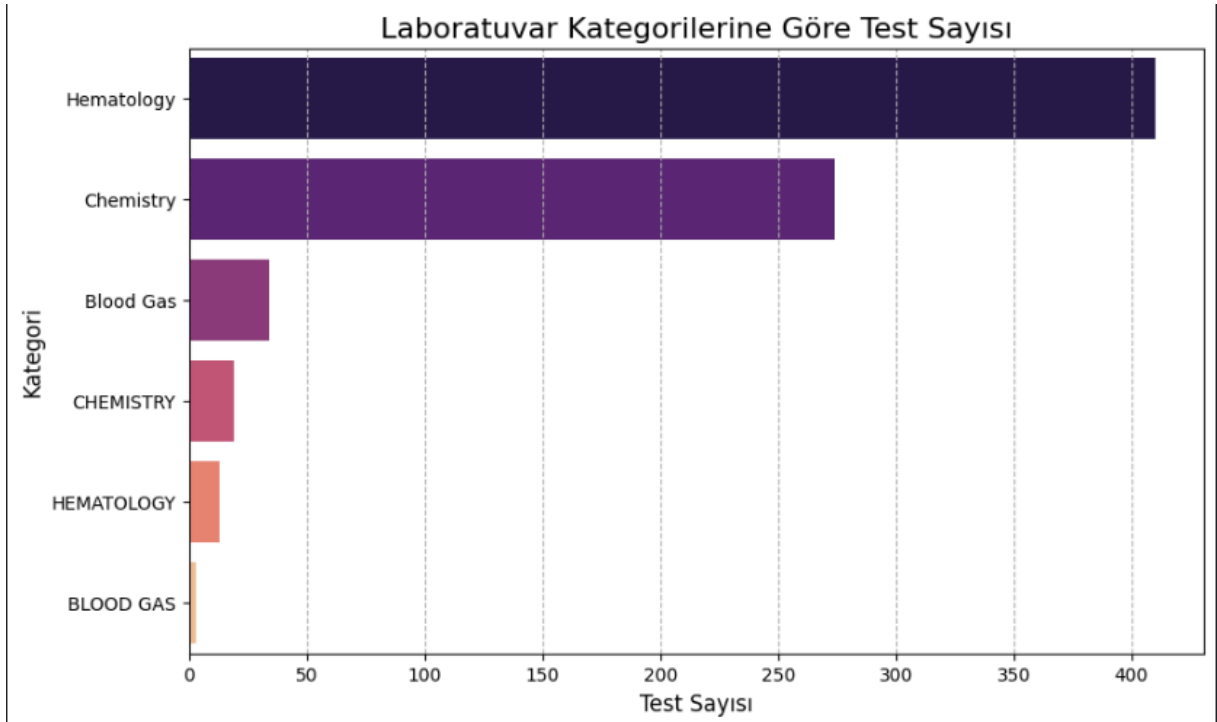
⚠️ **Doğrudan Model Girdisi Değildir** Bu tablodaki veriler (label , category vb.) doğrudan bir makine öğrenmesi modeline girdi olarak kullanılmaz. Bunun yerine, diğer tablolardan hangi sayısal verilerin çekileceğini ve nasıl işleneceğini belirlemek için bir **yol haritası** olarak kullanılır.

📌 **Sonuç** D_ITEMS.csv , MIMIC-III veritabanı ile çalışırken atılacak ilk ve en temel adımlardan birini oluşturur. Bu tabloyu doğru bir şekilde kullanarak, ham ve karmaşık olay verilerinden, sepsis tespiti gibi belirli bir klinik problem için hedefe yönelik, temiz ve standartlaştırılmış bir özellik seti oluşturabiliriz. Projenin başarısı, bu tablonun etkin bir şekilde kullanılmasına büyük ölçüde bağlıdır.

D_LABITEMS.csv

ID	Sütun Adı	Anlamı	🔗 Proje Bağlantısı
	itemid	Laboratuvar Testi ID'si	EN KRİTİK SÜTUN. LABEVENTS.csv 'deki itemid ile eşleştirilerek testin adını belirler.
	label	Test Etiketleri (Açıklama)	itemid 'nin insan tarafından okunabilir karşılığı (örn. "Lactate", "Creatinine"). Çok önemli.
	fluid	Vücut Sıvısı	Testin hangi sıvıdan yapıldığını belirtir (örn. "Blood", "Urine"). Sepsis için "Blood" testleri önceliklidir.
	category	Kategori	Testin ait olduğu laboratuvar bölümü (örn. "Chemistry", "Hematology", "Blood Gas"). Filtreleme için kullanışlıdır.
	loinc_code	LOINC Kodu	Testler için uluslararası standart kod. İleri analizler ve veri entegrasyonu için değerlidir.

row_id	itemid	label	fluid	category	loinc_code
0	1 50800	SPECIMEN TYPE	BLOOD	BLOOD GAS	NaN
1	2 50801	Alveolar-arterial Gradient	Blood	Blood Gas	19991-9
2	3 50802	Base Excess	Blood	Blood Gas	11555-0
3	4 50803	Calculated Bicarbonate, Whole Blood	Blood	Blood Gas	1959-6
4	5 50804	Calculated Total CO2	Blood	Blood Gas	34728-6



🧠 D_LABITEMS.csv : Kullanılabilirlik Özeti ve Proje İçin Rolü

🌟 **Genel Amaç** Bu tablo, LABEVENTS.csv tablosundaki sayısal verileri, klinik olarak anlamlı ve yorumlanabilir laboratuvar testi sonuçlarına dönüştüren bir **çeviri sözlüğüdür**. Bu tablo olmadan LABEVENTS tablosundaki verileri anlamlandırmak mümkün değildir.

✅ Model Hazırlığındaki Rolü

- **Özellik Seçimi:** Organ fonksiyon bozukluklarını (böbrek, karaciğer), pıhtılaşma sorunlarını ve enfeksiyonun ciddiyetini gösteren laboratuvar testlerinin **itemid** 'lerini belirlemek.
- **Veri Filtreleme:** Milyonlarca satırlık LABEVENTS tablosunu, sadece sepsis takibi için önemli olan test sonuçlarını içerecek şekilde filtrelemek.
- **Doğrulama:** Testlerin doğru vücut sıvısından (**fluid** = 'Blood') alındığını teyit etmek.

⚠️ **Doğrudan Model Girdisi Değildir** D_ITEMS.csv gibi, bu tablodaki veriler de doğrudan modele girdi olarak kullanılmaz. Bunun yerine, LABEVENTS.csv 'den hangi sayısal verilerin çekileceğini ve hangi klinik anlama geldiğini belirlemek için bir **yol haritası** olarak kullanılır.

📌 **Sonuç** D_LABITEMS.csv , sepsis modelimizin fizyolojik temelini oluşturan laboratuvar verilerini anlamlandırmamız için temel bir araçtır. Bu tablo sayesinde, sepsisin vücut üzerindeki etkilerini yansıtan en önemli biyobelirteçleri sistematik olarak belirleyip, analiz sürecimizi bu kritik verilere odaklayabiliriz. Bu, modelimizin doğruluğu ve klinik olarak anlamlılığı için hayati bir adımdır.

3.3. INPUTEVENTS_CV.csv & INPUTEVENTS_MV.csv: Tedavi Girdileri

Bu tablolar, hastaya uygulanan sıvı ve ilaç tedavilerini kaydeder. Hastanın durumundaki değişikliklerin hem nedeni hem de sonucu olabilecekleri için çok değerlidirler

row_id	subject_id	hadm_id	icustay_id	charttime	itemid	amount	amountuom	rate	rateuom	...	orderid	linkorderid	stopped	newbottle	originalamount	originalamountuom	originalroute	originalrate	originalrateuom	originalsite
0	1184	10114	167957	234989	2171-11-03 15:00:00	30056	400.0	ml	NaN	NaN	...	2557279	2557279	NaN	NaN	NaN	ml	Oral	NaN	NaN
1	1185	10114	167957	234989	2171-11-03 20:00:00	30056	120.0	ml	NaN	NaN	...	7828849	2557279	NaN	NaN	NaN	ml	Oral	NaN	NaN
2	1186	10114	167957	234989	2171-11-03 23:00:00	30056	120.0	ml	NaN	NaN	...	2744159	2557279	NaN	NaN	NaN	ml	Oral	NaN	NaN
3	1187	10114	167957	234989	2171-11-04 02:00:00	30056	120.0	ml	NaN	NaN	...	8475006	2557279	NaN	NaN	NaN	ml	Oral	NaN	NaN
4	1188	10114	167957	234989	2171-11-04 05:00:00	30056	120.0	ml	NaN	NaN	...	11183474	2557279	NaN	NaN	NaN	ml	Oral	NaN	NaN

--- Veri Tipleri ve Bellek Kullanımı ---

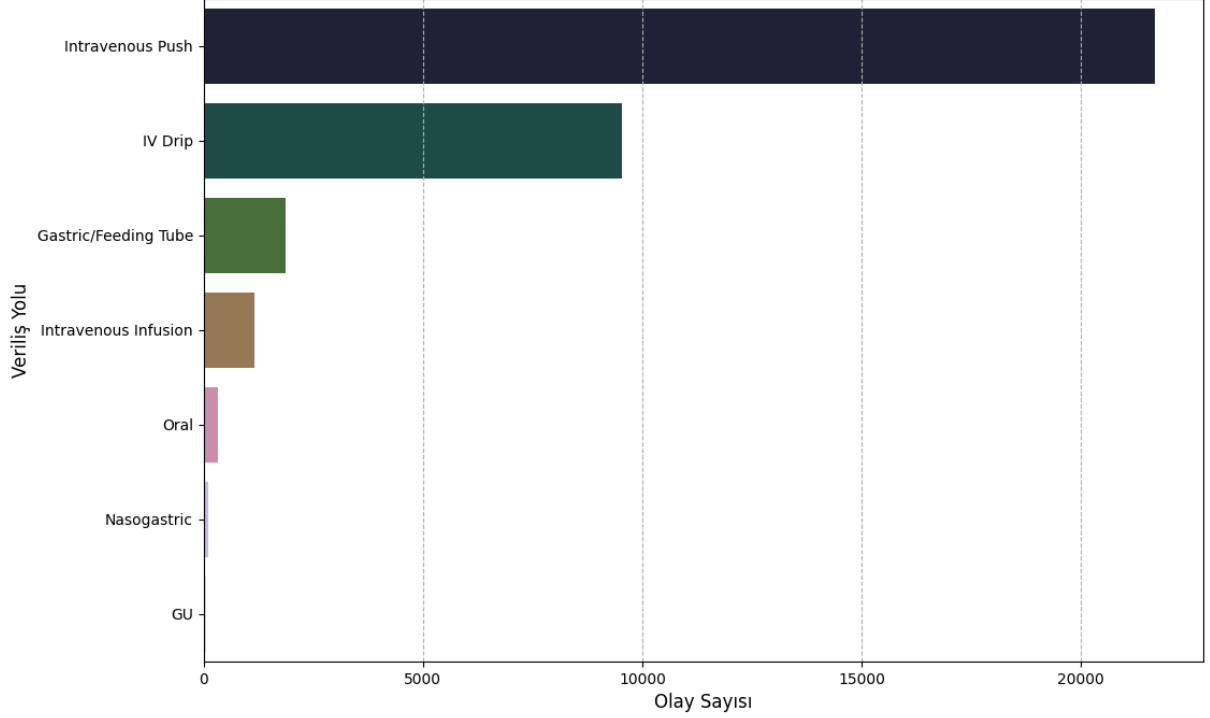
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 34799 entries, 0 to 34798
Data columns (total 22 columns):
#   Column                Non-Null Count  Dtype
---  ---
0   row_id                 34799 non-null  int64
1   subject_id             34799 non-null  int64
2   hadm_id                34799 non-null  int64
3   icustay_id             34799 non-null  int64
4   charttime              34799 non-null  datetime64[ns]
5   itemid                 34799 non-null  int64
6   amount                 24738 non-null  float64
7   amountuom              25003 non-null  object
8   rate                   9462 non-null   float64
9   rateuom                9548 non-null   object
10  storetime              34799 non-null  object
11  cgid                   34258 non-null  float64
12  orderid                34799 non-null  int64
13  linkorderid            34799 non-null  int64
14  stopped                810 non-null    object
15  newbottle              522 non-null    float64
16  originalamount          21214 non-null  float64
17  originalamountuom       25234 non-null  object
18  originalroute           34782 non-null  object
19  originalrate            11511 non-null  float64
20  originalrateuom         11515 non-null  object
21  originalsite            6 non-null     object
dtypes: datetime64[ns](1), float64(6), int64(7), object(8)
memory usage: 5.8+ MB
```

Sütunların Anlamı ve Proje İçin Önemi

Bu tablo, `metavision` versiyonuna göre daha basit bir yapıya sahiptir ancak yine de kritik bilgiler barındırır.

Sütun Adı	Anlamı	Proje Bağlantısı
<code>subject_id</code> , <code>hadm_id</code> , <code>icustay_id</code>	Hasta, Yatış ve YBÜ Kalış ID'leri	Diğer tablolara birleştirme için kritik anahtarlar .
<code>charttime</code>	Kayıt zamanı	Zaman serisi analizinde olayın gerçekleştiği zaman.
<code>itemid</code>	Verilen maddenin ID'si	<code>D_ITEMS.csv</code> ile birleştirilerek maddenin adı (örn. Normal Saline, Dopamine) bulunur. Çok önemli .
<code>amount</code>	Verilen miktar	Sıvı yüklemesi, ilaç dozu gibi hesaplamalar için temel sayısal değer.
<code>rate</code>	Verilme hızı	İnfüzyon hızını gösterir. Bu tabloda çok eksik veri olduğu için genellikle ihmal edilir.
<code>stopped</code>	Durdurulma durumu	Bir tedavinin sonlandırıldığını gösterir ('D/C'd' gibi).
<code>originalroute</code>	Veriliş yolu	Tedavinin nasıl verildiğini belirtir (Oral, IV gibi). Sepsis için "IV" olanlar kritiktir.

Girdi Olaylarının Veriliş Yoluna Göre Dağılımı (CareVue)



```
--- En Yaygın 5 Veriliş Yolu ---
originalroute
Intravenous Push      21695
IV Drip                9548
Gastric/Feeding Tube  1862
Intravenous Infusion  1167
Oral                   334
Name: count, dtype: int64
```


INPUTEVENTS_CV.csv : Kullanılabilirlik Özeti ve Proje İçin Rolü

🚩 **Genel Amaç** Bu tablo, eski `carevue` sisteminden gelen temel girdi verilerini içerir. Daha basit bir yapıya sahip olmasına rağmen, sepsis tedavisinin temel taşları olan **sıvı resüsitasyonu** ve **vazopresör kullanımı**ni tespit etmek için hayati önem taşır.

✅ Model Girdisi (Feature) Olarak Kullanılabilir Özellikler

📊 Özellik (Türetilmiş)	Kullanım Yeri	Açıklama
<code>hourly_iv_fluid_volume</code>	Sayısal giriş	<code>originalroute</code> 'u 'IV' olan ve <code>itemid</code> 'si temel sıvılara (örn. NaCl) denk gelen <code>amount</code> 'ların saatlik toplamı.
<code>is_on_vasopressor_hourly</code>	İkili (0/1) giriş	<code>itemid</code> 'si vazopresörlere (Norepinephrine, Dopamine vb.) denk gelen bir kaydın o saatte olup olmadığı.
<code>num_iv_meds_hourly</code>	Sayısal giriş	<code>originalroute</code> 'u 'IV' olan farklı ilaçların saatlik sayısı. Klinik aktivitenin bir göstergesi olarak kullanılabilir.

⚠️ Kullanılmaması Önerilen Alanlar

- `rate` ve ilgili sütunlar: Neredeyse tamamen boş.
- `originalamount`, `originalrate` gibi sütunlar: Tamamen boş.
- `cgid`, `orderid`, `linkorderid`: Klinik anlamdan çok, sistemsel ID'lerdir.

📌 **Sonuç** INPUTEVENTS_CV.csv , sepsis modellemesi için önemli bir veri kaynağıdır, ancak sınırlılıklarının farkında olunmalıdır. INPUTEVENTS_MV.csv ile birleştirildiğinde, hastanın tedavi geçmişinin tam bir resmini sunar. Özellikle sepsis için kritik olan **IV sıvı bolusları** ve **vazopresör ilaçların** varlığını tespit etmek için bu tablo taranmalıdır. `itemid` ve `originalroute` sütunları üzerinden yapılacak filtreleme, model için en değerli bilgileri ortaya çıkaracaktır.

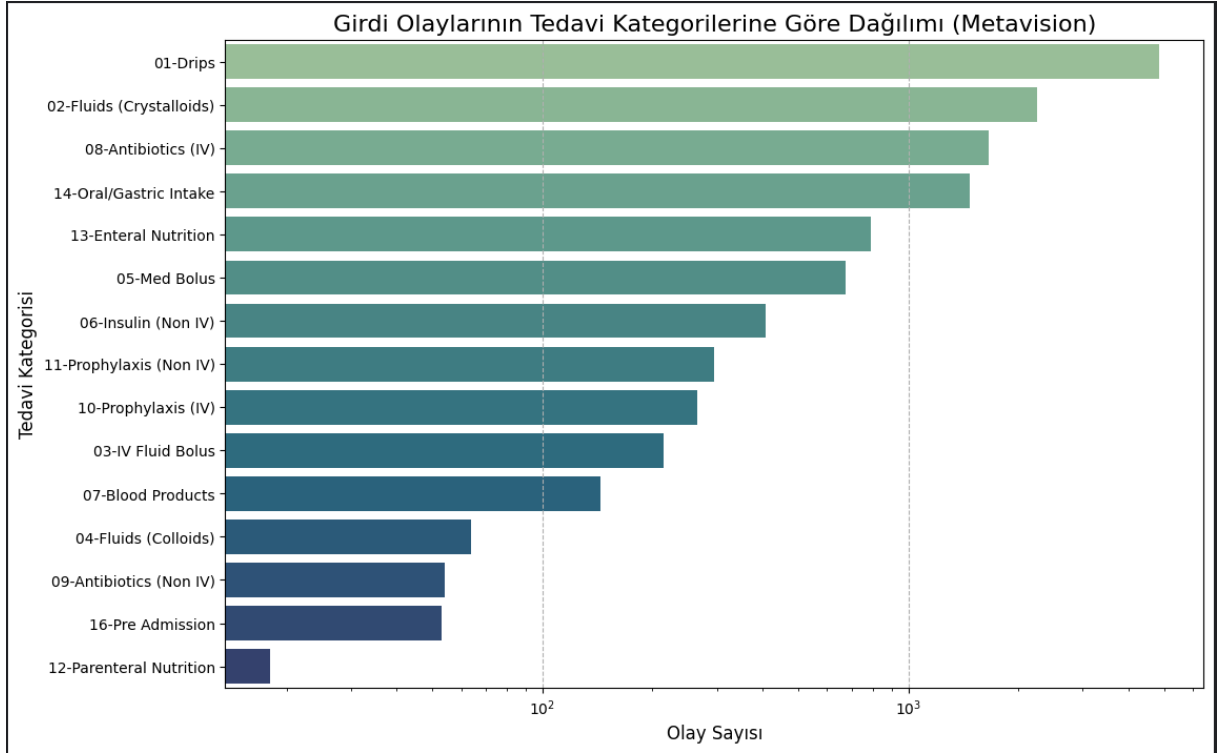
INPUTEVENTS_MV.csv

📊 Sütunların Anlamı ve Proje İçin Önemi

Bu tablo, tedavilerin detaylarını anlamak için çok zengin bilgiler içerir.

📊 Sütun Adı	📌 Anlamı	🔗 Proje Bağlantısı
<code>subject_id</code> , <code>hadm_id</code> , <code>icustay_id</code>	Hasta, Yatış ve YBÜ Kalış ID'leri	Diğer tablolarla birleştirme için kritik anahtarlar .
<code>starttime</code> , <code>endtime</code>	İnfüzyon başlangıç/bitiş zamanı	Olayın süresini net olarak verir. Zaman serisi için çok önemli .
<code>itemid</code>	Verilen maddenin ID'si	<code>D_ITEMS.csv</code> ile birleştirilerek maddenin adı bulunur.
<code>amount</code> , <code>rate</code>	Verilen miktar ve hız	Sıvı yüklemesi ve ilaç dozajını hesaplamak için temel değerler.
<code>ordercategoryname</code>	Tedavi kategorisi	Çok kullanışlı . <code>01-Drips</code> , <code>02-Fluids (Crystalloids)</code> gibi kategorilerle doğrudan filtreleme imkanı sağlar.
<code>ordercategorydescription</code>	Veriliş yöntemi	<code>Continuous IV</code> veya <code>Bolus</code> gibi bilgileri içerir, müdahalenin türünü anlamayı kolaylaştırır.
<code>patientweight</code>	Hasta kilosu (kg)	İlaç dozlarını kg başına standardize etmek için kritik .
<code>statusdescription</code>	Order durumu	Tedavinin neden durduğunu (<code>FinishedRunning</code> , <code>Rewritten</code> vb.) belirtir.

row_id	subject_id	hadm_id	icustay_id	starttime	endtime	itemid	amount	amountuom	rate	totalamountuom	isopenbag	continueinnextdept	cancelreason	statusdescription	comments_editedby	comments_cancelledby	comments_date	originalamount	originalrate	
0	118897	42367	130932	250305	2147-10-29 16:45:00	2147-10-29 22:57:09	60.000000	ml	NaN	—	ml	0	0	0	FinishedRunning	NaN	NaN	NaN	60.0	60.000000
1	118898	42367	130932	250305	2147-10-20 13:17:00	2147-10-20 23:25:58	10.000000	units	NaN	—	NaN	0	0	1	Rewritten	NaN	RN	2147-10-20 13:18:00	10.0	10.000000
2	118899	42367	130932	250305	2147-10-29 03:23:00	2147-10-29 03:53:00	99.999999	ml	199.999998	—	ml	0	0	0	FinishedRunning	NaN	NaN	NaN	100.0	200.000000
3	118900	42367	130932	250305	2147-10-22 22:00:00	2147-10-22 22:57:09	40.000000	ml	NaN	—	ml	0	0	0	FinishedRunning	NaN	NaN	NaN	40.0	40.000000
4	118901	42367	130932	250305	2147-10-16 06:21:00	2147-10-17 22:59:36	1309.899995	ml	54.999300	—	ml	0	0	0	FinishedRunning	NaN	NaN	NaN	1309.9	54.999298



--- En Yaygın 5 Tedavi Kategorisi ---

ordercategoryname	
01-Drips	4854
02-Fluids (Crystalloids)	2252
08-Antibiotics (IV)	1660
14-Oral/Gastric Intake	1477
13-Enteral Nutrition	789

✦ **Genel Amaç** Bu tablo, hastaya yapılan farmakolojik ve sıvı tedavisi müdahalelerini yüksek çözünürlükte anlamamızı sağlar. Bu müdahaleler, hastanın fizyolojik durumundaki değişikliklerin hem nedeni hem de sonucu olabilir, bu yüzden model için paha biçilmezdir.

✓ **Model Girdisi (Feature) Olarak Kullanılabilir Özellikler**

Öznitelik (Türetilmiş)	Kullanım Yeri	Açıklama
hourly_iv_fluid_volume	Sayısal giriş	ordercategoryname 'i 'Fluids' veya 'Blood Products' olan girdilerin saatlik toplam amount 'u.
is_on_vasopressor	İkili (0/1) giriş	itemid 'si herhangi bir vazopresöre ait bir kaydın o saatte (starttime ve endtime arasında) aktif olup olmadığı.
normalized_pressor_dose	Sayısal giriş	Vazopresörlerin rate değerlerinin patientweight 'e bölünerek standardize edilmiş hali (örn: mcg/kg/min).
had_iv_bolus_in_hour	İkili (0/1) giriş	O saat içinde ordercategorydescription 'ı 'Bolus' olan bir sıvı verilip verilmediği.

⚠ **Kullanılmaması Önerilen Alanlar**

- storetime, cgid, orderid, linkorderid gibi sistemsel ID'ler.
- comments... sütunları, serbest metin içerdiği için ilk aşamada modellemeye uygun değildir.

📌 **Sonuç** INPUTEVENTS_MV.csv, modern ve yapılandırılmış yapısıyla sepsis modellemesi için en değerli tablolardan biridir. Bu tablodan türetilen özellikler, modelin sadece hastanın anlık durumunu değil, aynı zamanda aldığı tedaviye verdiği yanıtı ve durumunun ciddiyetini de anlamasına yardımcı olacaktır. Projenin sonraki adımı, bu tablodaki bilgileri INPUTEVENTS_CV.csv ile birleştirerek ve diğer olay tablolarıyla (CHARTEVENTS , LABEVENTS) senkronize ederek kapsamlı bir özellik matrisi oluşturmaktır.

5. Sonuç ve Sonraki Adımlar

Bu keşifsel veri analizi, MIMIC-III demo veri setinin, sepsis tespiti için zengin ve anlamlı bilgiler içerdiğini göstermiştir. Veri tabloları arasındaki ilişkiler, eksik veri yapıları ve sepsis için kritik olan klinik belirteçler başarıyla tanımlanmıştır.

Sonraki Adımlar:

1. Yukarıda belirtilen yol haritasına göre veri işleme ve birleştirme adımları kodlanacaktır.
2. Özellik mühendisliği ve etiketleme süreci tamamlanacaktır.
3. Baseline model eğitilerek ilk sonuçlar alınacak ve model performansı değerlendirilecektir.
4. Elde edilen bulgulara göre LSTM modeli geliştirilip sonuçlar karşılaştırılacaktır.