

# **Predictive Fleet Analytics**

## **方法论**

# Predictive Fleet Analytics 方法论

## 产品概述

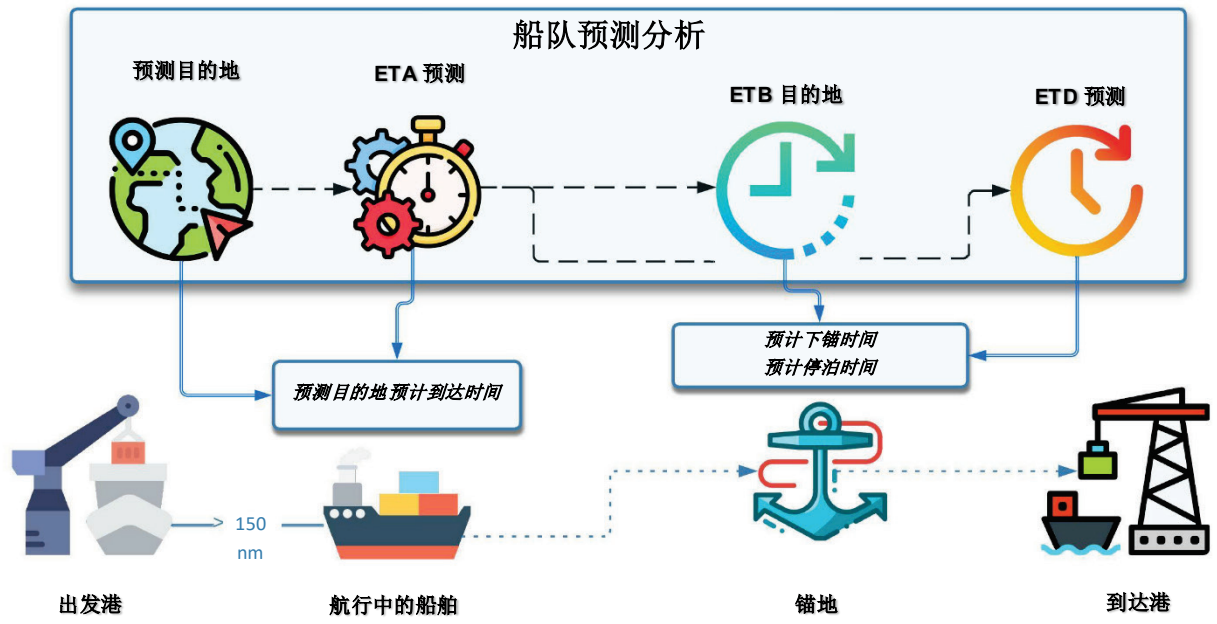
Lloyd's List Intelligence 的船队预测分析系统将我们自有和可信的船舶及其动态基础数据与先进的机器学习技术相结合，可让客户更准确地了解感兴趣船舶最为可能的目的地、可能的到达时间，以及港口拥塞状况如何影响其停泊和离港时间。

该服务是应客户需求而创建的，客户明确要求获得比标准 AIS 数据提供的视图更可靠和更完整的未来船舶动态视图。“目的地”和“预计到达时间”(ETA)信息由船员手动输入到 AIS 传输系统中，通常不准确。这会对客户产生许多连锁影响，比如无法主动规划业务运营、无法预测可能存在合规或安全风险的船舶何时到达感兴趣的海域等。

船队预测分析系统为用户提供了更可靠的决策基础。它汇集了感兴趣船舶的前瞻性信息（有关预测的目的地和预测的 ETA），以及有关港口拥塞级别、当前和历史贸易路线趋势的补充见解。

通过此工具，用户可以了解以下信息：

- 船舶的目的地 - *预测的目的地*
- 可能的到达时间 - *预测的 ETA 时间*
- 船舶到达泊位的可能时间 - *预测的“预计停泊时间”(ETB) - 考虑了该船舶类型在该港口的拥塞和等待时间*
- 可能的船舶离港时间 - *预测的“预计离港时间”(ETD) - 考虑了该船舶类型在该港口的周转时间*
- 有关感兴趣地点的历史拥塞数据和拥塞预测 - *基于未来 5 天预计将到达该地点的所有船舶数*
- 当前在感兴趣的贸易路线上航行的船舶，以及过去曾在该路线上航行的船舶，*这方面信息通过贸易路线进行分析*



PFA 提供的有关船舶航行的预测信息的简要概述

# 模型协同

简要概述了我们的三种预测模型如何协同工作

## PFA 模型

船队预测分析系统使用了三种预测模型：

1. 预测目的地
2. 预计到达时间
3. 港口拥塞情况

通过这些模型，我们可以深入了解下一个预测目的地、预计到达时间、预计停泊时间、预计离港时间，以及船舶交易和航行信息。这些信息在工具中是动态变化的，因此，用户可以设置在感兴趣船舶的预测数据发生变化时发出警报（*预期 2022 年第 3 季度推出此功能*）。

## 模型工作原理

三种预测模型遵循不同的建模流程，但同时也协同工作，以提供准确且及时的船舶航行预测见解。

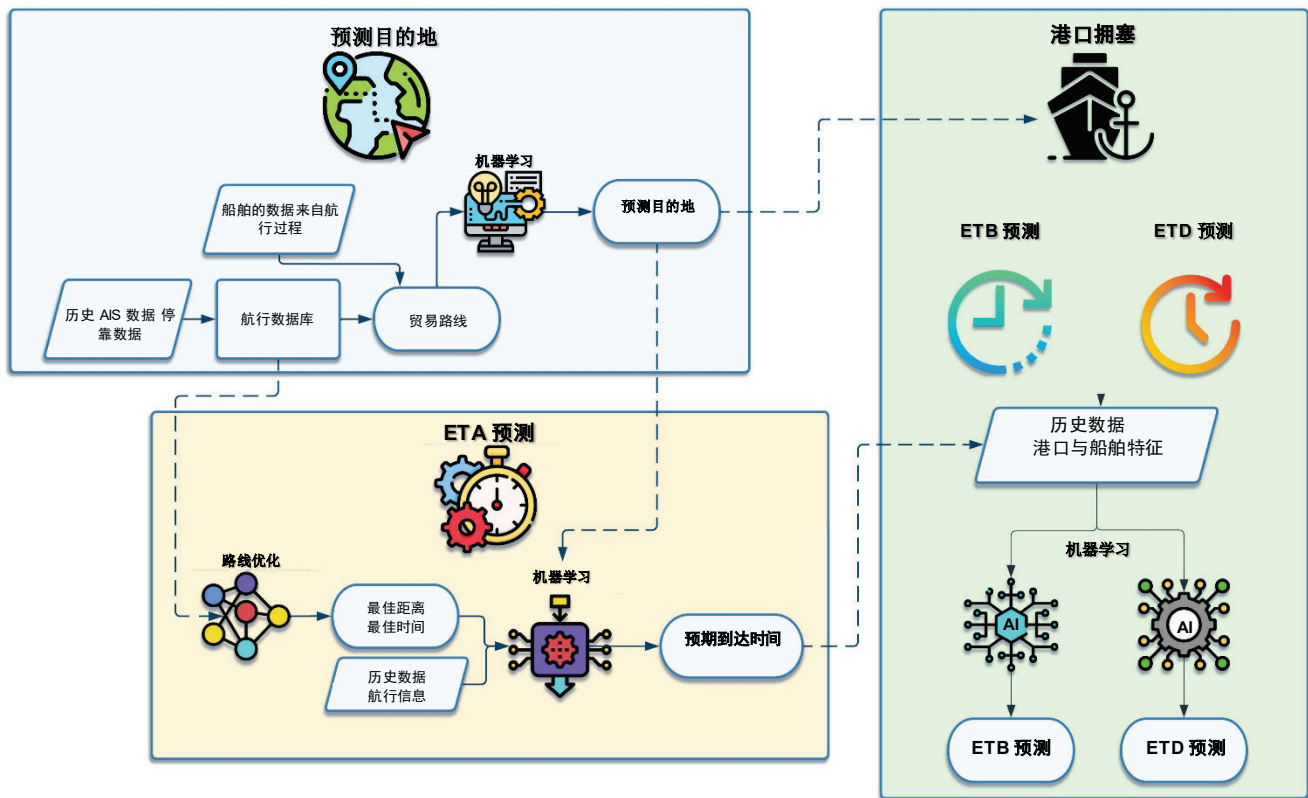
船舶从前一个港口出发航行约 150 海里后，预测目的地模型将使用历史和实时船舶航行数据以及机器学习来预测最为可能的下一个目的地。该模型创建了前三个最为可能的目的地（港口）列表以及到达每个目的地的概率。每个目的地的到达概率分别进行评分，总分为 100%。概率最高的目的地始终显示在第一位。

生成下一个预测目的地后，预计到达时间 (ETA) 机器学习模型会预测到达预测目的地的时间，该过程使用了源于我们历史航行数据库的数据点网络信息（最重要的是该类型和尺寸的船舶经过预测距离的航行到达目的地所花的平均时间）。预测距离源于我们的历史航行数据库，应该很好地反映了我们预期该船舶将采用的真实路线。

接下来，港口拥塞模型会使用“预测目的地”和“预计到达时间”来为船舶进一步预测“预计停泊时间” (ETB) 和“预计离港时间” (ETD)。这些预测源于我们广泛的港口层级数据，通过这些数据，我们能够在泊位层面了解和掌握实时和历史的船舶动态（包括等待和周转时间）。

我们的所有模型都为该预测流程提供了以下基本输入项：

- “预测目的地”是一个基本输入项，因为 ETB 和 ETD 预测是特定于港口的；
- “预测 ETA”时间至关重要，因为预测过程还与日期相关；
- 最后，我们值得信赖的船舶特征数据库十分关键，因为有关拥塞和等待时间的预测是针对特定船舶类型和尺寸提供的。



更精细地了解 PFA 模型及其协同效应

我们使用的模型是动态的，它会定期更新，以便根据最新数据持续验证预测的准确性。船舶在其航程中再往前航行约 30 海里后，模型会重新计算，如果新的预测与以前的预测不同，则会重新公布预测。

## 等待时间智能化

根据预测的 ETA、ETB 和 ETD 数据，可以计算出在锚点等待的预估时间长度以及船舶在港口将花费的预估时间长度。

- 对于预测等待时间 - 通过计算 ETA（船舶到达锚点时）与 ETB 之间的差值来完成。
- 对于预测在港时间 - 计算 ETB 和 ETD 之间的差值。

此外，还可从该模型创建船舶级交易模式信息和当前贸易路线分析。船舶的贸易路线被认为是预测目的地模型的建模步骤之一。反过来，这让我们能够提供“贸易路线”分析，让用户能够分析当前位于感兴趣贸易路线上的船舶，以及之前在规定时间内穿越感兴趣贸易路线的船舶。

历史和预测港口拥堵等级是按 5 天的时间范围提供的，并由港口拥堵模型提供技术支持。这些数据基于该船舶尺寸和类型的历史拥堵状况，以及根据目前位于港口中以及预测很快将到达该港口的类似船舶的预测拥堵状况。



# 预测流程

预测的阈值是什么，我们如何生成这些阈值？

## 对船舶进行预测的要求.....

我们已将阈值和考虑因素构建到船队预测分析模型中，以确保我们的预测数据尽可能可靠且信息量充足。请注意，该工具中的三个模型（预测目的地、预计到达时间和港口拥塞情况）协同工作、相互依赖，以便进行准确和及时的预测。因此，在不满足某些条件的情况下，不会显示船舶的预测信息。

我们的预测流程通过以下步骤进行：

- 首先，在考虑任何港口、子港口或停靠锚点的同时，将检索船舶的 AIS 位置并清除其中的任何无效和（或）缺失数据。当为船舶生成新位置时，每 6 分钟重复一次此过程。
- 其次，根据船舶的位置和停靠状况，将以下状态之一分配给船舶：
  - 正在航行 - 表示船舶正在航行
  - 在锚点区 - 表示船舶停在锚点区
  - 驶向港口 - 表示船舶正在向港口或泊位移动
  - 在港口 - 表示船舶位于港口内或泊位处
- 第三，如果满足以下三个条件，则会预测船舶的目的地：
  - 船舶正在航行
  - 船舶自最后一个目的地出发以来至少航行了 150 海里
  - *如果我们提供了一个更新的预测*，即预测船舶的前一个目的地距离船舶当前位置至少 30 海里

最后，生成预测目的地后，便可以生成路线优化和 ETA 预测。反过来，这些数据允许模型生成 ETB 和 ETD 预测。

# “预测信息不可用”

## 预测信息不可用的原因

主要原因是：

- 船舶未传输 AIS 数据或 AIS 数据无效
- 船舶的 AIS 位置和停靠状况不会将船舶的状态分类为“正在航行”
- 自前一个港口出发起，船舶还需至少航行 150 海里



# 船舶操作状态

## 定义船舶在航行期间的操作状态

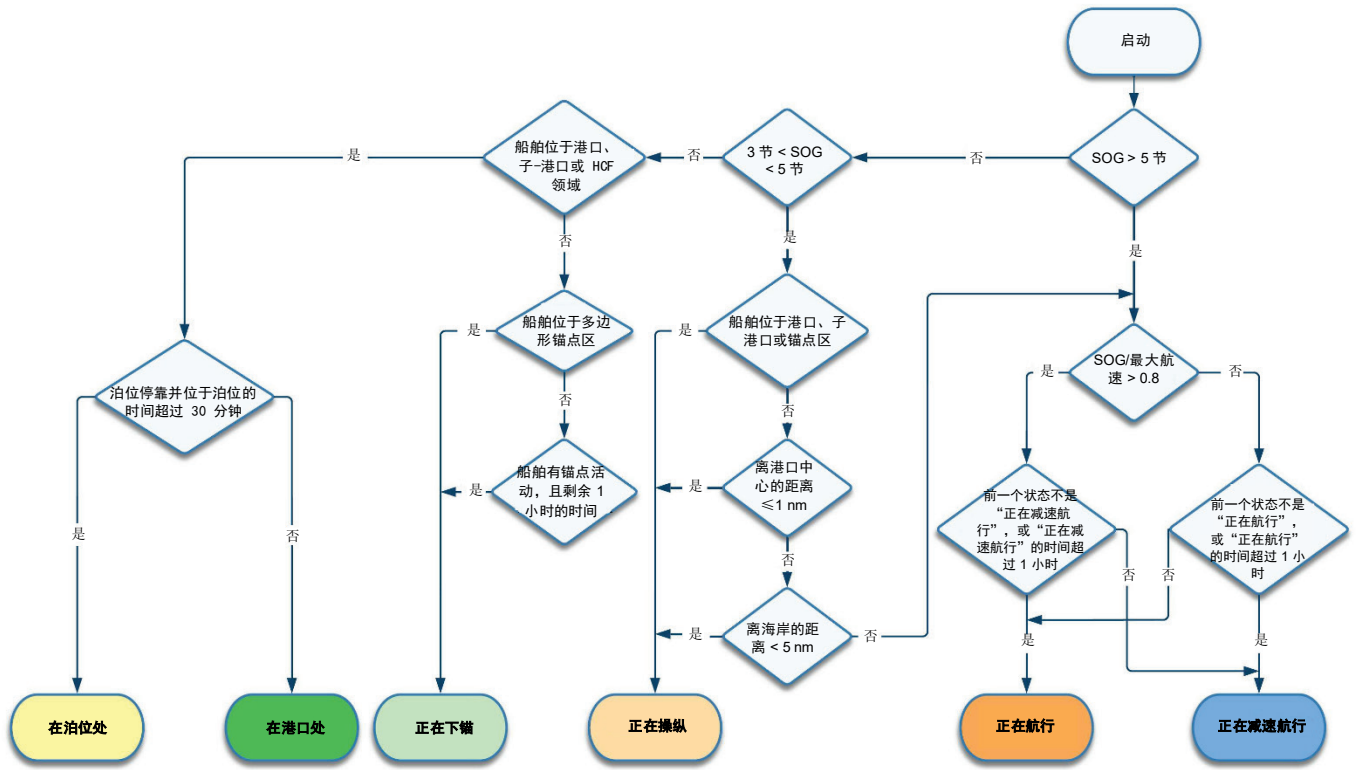
船舶的操作状态提供了关于船舶航行期间当前位置和状态的额外信息。

操作状态由 AIS 传输数据确定，其中包括船舶的航速、吃水深度和尺寸。每次收到新的 AIS 消息时，我们都会使用许多内部行为参数来更新船舶的操作状态。

下表列出了所有可能的操作状态：

操作状态	描述	活动
<i>在泊位处</i>	船舶已在泊位内下锚	已停止
<i>在港口处</i>	船舶位于港口内	已停止
<i>正在下锚</i>	船舶正在锚点区内下锚，或正在离任何港口中心超过 5 nm 的地方缓慢移动	已停止
<i>正在操纵</i>	船舶正在操控。与巡逻非常相似，但发生在靠近海岸或港口区域的地方，或在离港口中心不到 5 nm 的地方缓慢移动	正在航行
<i>正在减速航行</i>	船舶正在海上减速航行	正在航行
<i>正在航行</i>	船舶正在加速或高速运行	正在航行

下面的流程图显示了我们用于确定船舶操作状态的逻辑，具体取决于最新 AIS 传输系统中的字段满足的标准：



船舶操作状态流程

# 海域状态

## 定义船舶在航行期间的海域位置

船舶的海域状态更详细地说明了在全球海域内花费的时间和航行的距离，从而更好地了解随着时间的推移船舶的使用情况和交易模式。

船舶的海域状态是根据 LLI 的 AIS 偏航逻辑通过船舶的 AIS 传输系统来确定的，从而降低了 AIS 位置错误的可能性，并仅提供相关的 AIS 消息。

然后，这些 AIS 位置根据 [www.marineregions.org](http://www.marineregions.org) 网站上的“海洋边界”进行分类，该网站提供了查看船舶作业区域的标准化方式。《联合国海洋法公约》(UNCLOS) 规定中包含的所有区域都包含在 Seasearcher 中。下表列出了这些区域和其他区域；

区域名称	描述
领海 (TW)	从沿海国基线向外延伸最多 12 NM 的一带沿海水域。这也包括上面的空域和下面的海床。
内海 (IW)	群岛国国家领海 (TW) 向陆一侧的水域，其中包括河流和运河
毗连区 (CZ)	从领海向外延伸的一带水域 (TW 至 24 NM)，所属国家可在其中对海洋作业实施管控，以防止或惩罚违反其法律的行为。
群岛水域 (AW)	根据 UNCLOS 第 47 条划定的群岛国群岛基线所封闭的水域
公海	不属于任何国家管辖范围的开放海洋水域
排放控制区 (ECA)	为最大限度减少 1997 年《防止船舶污染国际公约》附件 VI 所界定的船舶气载排放物而实施更严格的管制的海洋区域。
国际航道测量组织 (IHO)	IHO 海洋区域由 <a href="#">国际航道测量组织</a> 确定。由 IHO 于 1953 年出版的《海洋边界》(第 23 号特种出版物) 组成。它代表了世界主要海洋的边界。
联合战争委员会区域 (JWC)	由联合战争委员会 (JWC) 指定的、会对保险公司造成更大风险的区域。
受制裁的专属经济区 (受制裁的 EEZ)	当前受到来自以下机构制裁影响的国家的专属经济区：美国财政部外国资产控制办公室 (OFAC)、欧盟 (EU)、英国财政部 (HM Treasury)、联合国 (UN)
非受制裁的专属经济区 (非受制裁的 EEZ)	自一个国家的海岸线/海岸边界向外延伸 200 NM 的非受制裁海域，所属海岸国对其内的所有经济资源拥有管辖权

您可以在下面通过 CSV 下载完整的海域列表，同时可以在 Seasearcher 的“位置”页面上查看我们的高风险海域、ECA、领海和 EEZ。如果船舶同时位于多个海洋区域（例如俄罗斯 JWC 和波罗的海 ECA）内，则将这些区域归类为这种海域。



Comprehensive Maritime Regions.xlsx 2MB

二进制

当船舶进入给定区域并对位置进行分类时，我们会收集以下指标：

- 船舶在该区域内第一条 AIS 消息和最后一条 AIS 消息的时间戳
- 在该区域内航行的距离 (nm)
- 在该区域内的总时间

# PFA 支持的船舶类型和尺寸

# 船舶类型

## 船队预测分析支持的所有船舶类型

船队预测分析涵盖了多种船舶类型。下面按字母顺序列出这些类型。

如果您正在查询的船舶不是列出的以下 LLI 船舶类型之一，则它未包括在船队预测分析中。

## LLI 船舶类型

运酸船

沥青运输船

散装集料船

散装货船

集装箱容量散装货船

散装水泥船

散装矿砂船

供油船

化学品运输船

散货和石油两用船

化学品和石油两用船

液化石油气和液化天然气两用运输船

矿油两用船

压缩天然气运输船

原油运输船

食用油运输船

渡船

鱼油船

浮式钻井生产船

果汁运输船

全蜂窝式集装箱船

全格舱式冷藏船

普通货船

---

集装箱容量普通货船

---

液化天然气运输船

---

液化石油气运输船

---

牲畜船

---



糖浆运输船

---

海军辅助舰船

---

客船（游轮）

---

客滚船

---

客船（未指明）

---

成品油船

---

冷藏船

---

滚装船

---

集装箱容量滚装船

---

供应船

---

油船（未指定）

---

油轮和集装箱运输船

---

适合公路运油车的滚装船

---

车辆运载船

---

运水船

---

酒类运输船

---

木片运输船

---

# 船舶尺寸

## 每种船舶类型的尺寸

PFA 支持每种船舶类型的多个不同尺寸，例如散装货船的灵便型、大灵便型和超灵便型，或者油轮的巴拿马型、阿芙拉型和苏伊士型。这些区别是根据客户反馈实施的，以使用户能够尽快细分到感兴趣的船舶类型和尺寸。

每种船型类别中的船舶尺寸细目如下表所示：

船舶类型	船舶尺寸
散装货船	小型, 0'-10' dwt
散装货船	灵便型, 10'-35' dwt
散装货船	大灵便型, 35'-50' dwt
散装货船	超灵便型, 50'-60' dwt
散装货船	巴拿马型, 60'-100' dwt
散装货船	海岬型, 100'-200' dwt
散装货船	超大型散货船, 200'+ dwt
普通货船	0'-5' dwt
普通货船	5'-10' dwt
普通货船	10'-20' dwt
普通货船	20'+ dwt
冷藏散装船	0'-2' dwt
冷藏散装船	2'-6' dwt
冷藏散装船	6'-10' dwt
冷藏散装船	10'+ dwt
集装箱船	小支线船, 0'-1' TEU
集装箱船	支线船, 1'-2' TEU
集装箱船	大支线船, 2'-3' TEU
集装箱船	巴拿马型, 3'-5' TEU
集装箱船	超巴拿马型, 5'-8' TEU
集装箱船	新巴拿马型, 8'-12' TEU
集装箱船	巨型集装箱船, 12'-14.5' TEU
集装箱船	超大型集装箱船, 14.5'+ TEU
滚装船	0'-5' dwt
滚装船	5'-10' dwt
滚装船	10'-15' dwt

滚装船	15'+ dwt
辆运载船	0'-30' gt
车辆运载船	30'-50' gt
车辆运载船	50'+ gt
化学品运输船	0'-5' dwt
化学品运输船	5'-10' dwt
化学品运输船	10'-20' dwt
化学品运输船	20'+ dwt
油轮	成品油, 0'-5' dwt
油轮	成品油, 5'-10' dwt
油轮	通用型, 10'-25' dwt
油轮	MR 1, 25'-35' dwt
油轮	MR 2, 35'-45' dwt
油轮	巴拿马型 (LR 1), 45'-80' dwt
油轮	阿芙拉型 (LR 2), 80'-120' dwt
油轮	苏伊士型, 120'-200' dwt
油轮	巨型油轮, 200'-320' dwt
油轮	超大型油轮, 320'+ dwt
液化气体运输船	液化石油气, 0'-50' cbm
液化气体运输船	液化石油气巨型气体运输船, 50'+ cbm
液化气体运输船	液化天然气, -100' cbm
液化气体运输船	液化天然气巨型气体运输船, 100'-200' cbm
液化气体运输船	液化天然气卡塔尔型, 200'+ cbm
其他液体运输船	-1' dwt
其他液体运输船	1'+ dwt
游轮	0'-2' gt
游轮	2'-10' gt

游轮	10'-60' gt
游轮	60'-100' gt
游轮	100'-150' gt
游轮	150'+ gt
渡船 - 客货滚装	0'-2' gt
渡船 - 客货滚装	2'-5' gt
渡船 - 客货滚装	5'-10' gt
渡船 - 客货滚装	10'-20' gt
渡船 - 客货滚装	20'+ gt
渡船 - 仅限客船	0-299 gt
渡船 - 仅限客船	300-1' gt
渡船 - 仅限客船	1'-2' gt
渡船 - 仅限客船	2'+ gt

# 预测目的地模型

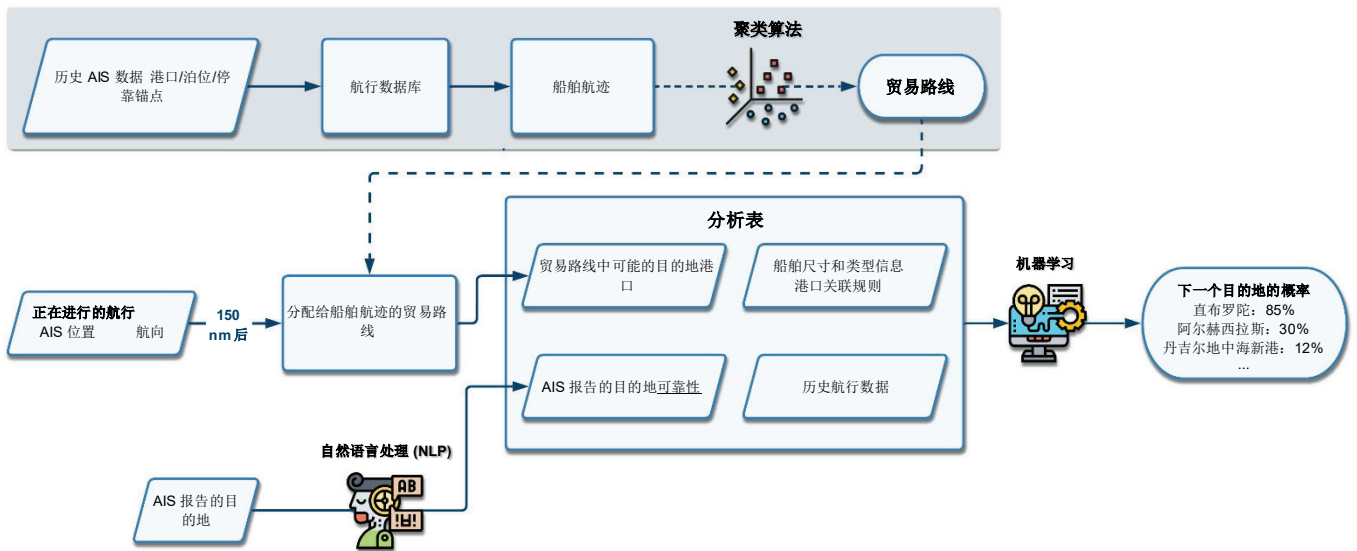
# 概述

## 预测目的地模型的简要概述

预测目的地模型力求在每艘船舶离开前一个港口并航行约 150 海里后确定其最为可能的目的地。

在将任何数据输入模型之前，需要完成一些初始步骤。第一步是创建“航行数据库” - 一个“干净”航行的历史数据集，源于我们的历史动态数据库。在此基础上，该模型能够通过将类似的航迹合并在一起定义船舶航迹（或贸易路线）组。这些贸易路线随后由该模型用于根据船舶最新的航迹确定其最为可能的港口。

然后，该模型会使用其他因素，例如我们预期从该尺寸和类型的船舶所看到的行为，该船舶是否与小规模港口分组相关联，以及船舶在声明正确目的地或 ETA 方面的可靠性。这与贸易路线和航迹结合使用，以进一步验证该船舶的可能目的地。



预测目的地模型摘要

# 航行数据库

什么是“航行数据库”？

我们根据过去三年的航行数据（可追溯至 2019 年 1 月），创建了一个被称为“航行数据库”的数据集。

此数据集包括数百万个船舶位置数据点。这些位置被排列到每艘船的航程中。单次航行定义为船舶的港口到港口（有时为子港口或锚点区）连线。然后使用一个“干净”历史航行的大型子集来计算船舶航迹，贸易航线则是基于船舶航迹构成的。

此过程是预测目的地模型的基本步骤。“航行数据库”利用我们自己的历史数据集，我们直接拥有和维护这些数据集，以确保最高水平的覆盖范围、及时性和准确性。

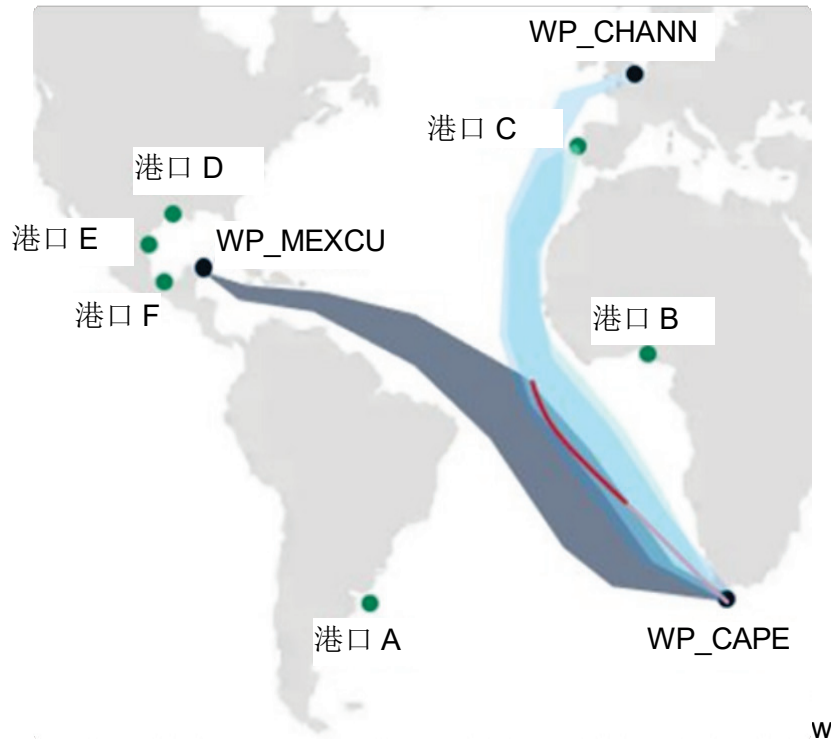


# 贸易路线

什么是“贸易路线”？

“预测目的地”模型的一个组成部分是识别船舶正在航行的贸易路线。相互“接近”的（即在整个或部分航程所遵循的路线中是相似的）历史船舶航迹形成一条贸易路线。

贸易路线为模型提供了重要信息，因为它们大大限制了船舶可能前往的港口数量。



正在进行的船舶航迹（红色）和三组历史航迹（或贸易路线）（蓝色）的示例

为了形成贸易路线，我们利用了历史船舶航迹。航迹定义为跨越 150 海里的一系列位置，也考虑了航向。

然后将 K 均值（一种无人监督的机器学习算法）用于船舶航迹，以评估它们之间的相似性并创建群集。

根据船舶在当前航程中的位置和航向，计算一条航迹，并为船舶分配一组航迹（或贸易路线）。指定的贸易路线是与船舶航迹最为相似的贸易路线。这有助于限制船舶的可能目的地数量。

船舶航行至少 30 海里后，将重复此过程，如果生成了新的预测或船舶更改其 AIS 传输系统中报告的目的地，用户将收到通知。

# 机器学习算法

我们如何在预测目的地模型中利用机器学习？

在由贸易路线识别流程所提供的信息与船舶和历史航行数据涉及的其他影响特征信息之间，机器学习模型用于建立进一步的预测联系。

使用**逻辑回归**模型产生特定船舶驶往特定港口的概率。为了更为明确，模型提出以下问题：

- “船舶的下一个目的地是 ‘X’ 港口的可能性有多大？”，其中“X”是船舶贸易路线内的一个港口。

此程序针对船舶当前贸易路线内的所有港口重复执行。为当前正在行驶的船舶分配一条贸易路线是一个关键步骤，因为它允许模型限制需要计算概率的港口数量。

根据上述过程，模型输出项是根据船舶驶往其中每一个港口的概率排列的港口列表。因此，可以推导出船舶的最高概率目的地，在某些情况下，也可以推导出概率第二和第三的目的地。

**请注意**，每个潜在目的地均按其自身指标进行评分，评分范围为 0-100%。

# 模型特征

## 在预测目的地模型中输入的特征

该模型使用了多个特征作为输入项，以帮助准确预测船舶的下一个目的地。这些特征包括船舶出发港、预测港口和船舶尺寸或类型、交易模式和 AIS 目的地可靠性指标等关联规则，其中可靠性指标根据历史数据中存在的船舶报告目的地的准确性计算得到。

下表提供了模型特征的细化视图以及简短描述：

特征名称	取值	描述
<i>同区船舶</i>	是/否	指示船舶是否按月在同一专属经济区 (EEZ) 中航行
<i>同港口船舶</i>	是/否	指示船舶过往是否停靠同一港口
<i>同一港口</i>	是/否	指示预测的港口是否是船舶最近的历史数据中经常停靠的港口
<i>区域港口</i>	是/否	指示预测的港口是否在船舶过往曾在其中航行的 EEZ 之内
<i>AIS 目的地可靠性</i>	数字	AIS 报告目的地的可靠性指标
<i>按船舶尺寸的相对频率</i>	数字	按船舶尺寸分类，预测港口的历史停靠相对频率
<i>航程计数</i>	数字	到达预测港口的历史行程数
<i>港口停靠的相对频率</i>	数字	特定船舶曾经停靠预测港口的相对频率
<i>港口关联规则</i>	数字	出发港和到达港之间过往的连线有多少次？
<i>按船舶类型的相对频率</i>	数字	按船舶类型分类，预测港口的历史停靠相对频率
<i>按船舶大小的关联规则</i>	数字	出发港、到达港和船舶尺寸之间的关联规则
<i>按船舶类型的关联规则</i>	数字	出发港、到达港和船舶类型之间的关联规则
<i>预测港口的相对距离</i>	数字	离预测目的地的优化距离与船舶已航行的历史平均距离之间的比较

# 特征重要性

## 预测目的地模型的特征重要性

特征重要性是指模型中为每个特征所赋予的权重，以进行准确的预测。当模型经常依赖模型特征进行预测时，模型特征就更为重要。

下表从影响最大到最小列出了预测目的地模型的八个特征：

排名	特征名称	描述
1	<i>同一港口</i>	指示预测的港口是否是船舶最近的历史数据中经常停靠的港口
2	<i>同区船舶</i>	指示船舶是否按月在同一专属经济区 (EEZ) 中航行
3	<i>区域港口</i>	指示预测的港口是否在船舶过往曾在其中航行的 EEZ 之内
4	<i>同港口船舶</i>	指示船舶过往是否停靠同一港口
5	<i>AIS 目的地可靠性</i>	AIS 报告目的地的可靠性指标
6	<i>按船舶尺寸的相对频率</i>	按船舶尺寸分类，预测港口的历史停靠相对频率
7	<i>航程计数</i>	到达预测港口的历史行程数
8	<i>港口停靠的相对频率</i>	特定船舶曾经停靠预测港口的相对频率

# AIS 报告的目的地匹配

## 我们如何利用船舶所报告的目的地？

在确定船舶的下一个目的地时，AIS 消息中报告的目的地通常非常不准确和不可靠。我们的客户会定期向我们报告与此相关的这些挑战。作为我们于 2021 年对原始 AIS 数据进行的内部研究的一部分，这些数据也得到了加强 - 研究发现自航程结束后 5 天内，不到 30% 的船舶报告了正确的或可清晰识别的目的地（可随时与我们的港口和地点数据库匹配）。

但是，在某些情况下，它仍然包含有价值的信息，可在预测目的地模型中予以提取和使用。

为了最大限度地利用这些数据，我们引入了一个预处理步骤。这会使用[自然语言处理](#)技术来处理原始 AIS 中经常出现的模糊或错误的目的地拼写变体，将其与我们的港口和地点数据库中最接近的实用已知位置（如果有的话）进行匹配。

确定精确匹配或最接近匹配后，将根据特定船舶 AIS 报告目的地的历史可靠性进行加权。这将在以后用作机器学习模型的输入项。

# 预计到达时间模型

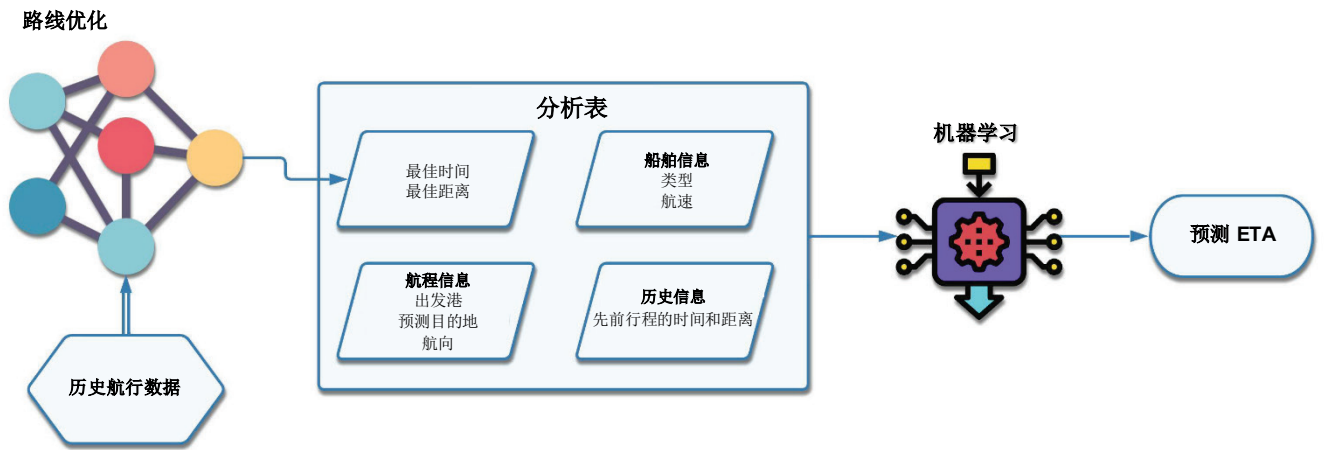
# 概述

## ETA 模型的简要概述

此模型可预测船舶到达的时间 - 船舶将到达其预测目的地的时间。“预测 ETA”旨在准确判断船舶何时可能到达预测目的地的锚点区。

ETA 预测模型分为以下两个步骤，本节将详细介绍它们：

1. 路线优化
2. 机器学习模型



ETA 预测模型概述



# 路线优化

路线优化根据以下 3 个关键因素为船舶推导出要遵循的最佳路线：

- 出发港
- 当前位置
- 预测目的地

根据航行数据库中船舶过往最常采用的路线，创建到达预测目的地的最佳路线，让我们能够估计船舶完成航行所需的剩余距离。

除此之外，我们还考虑了船舶类型和该船舶预期的平均航速，以使模型能够准确估计完成当前航行的最佳时间。

# 我们如何创建优化的路线？

我们已使用人工神经网络来推导船舶到达预测目的地的预期距离和时间。

为此，我们使用神经网络来了解从一个港口航行到另一个港口时，该船舶类型和尺寸所采用的最常见路线。此神经网络是通过一系列经过验证的先前船舶航迹来创建的。

该模型使用了最短路径算法（[Dijkstra 算法](#)），该算法根据观察到的类似类型和尺寸的船舶的历史行为确定船舶的最佳路线。通过确定该船舶类型的可能航速，该模型能够计算船舶到达预测目的地的平均时间和距离。

此外，我们还利用[梯度提升](#)进一步帮助验证我们对预计到达时间的预测。该算法可进一步帮助我们识别模型中的任何缺失数据或离群值，以便相应地解决和适应这些数据或离群值。此机器学习算法已在文献中用于实时预测和一般的 ETA 预测[1、2]，特别是在海事领域[3]。

# 模型特征

## 用作 ETA 模型输入的特征

模型的特征（或输入项）包括船舶当前航程的起点、到达其预测目的地的最佳时间和距离、船舶的特定特征（如船舶的类型和大小）、当前航速和航向以及基于船舶先前航行的历史信息。

下表提供了模型特征的细化视图以及简短描述：

特征名称	取值	描述
<i>到达港</i>	字符	预测目的地模型中的预测目的地
<i>出发港</i>	字符	开始航行的港口的名称
<i>平均航速</i>	整数	过去 30 NM 的平均对地航速
<i>平均航程距离</i>	整数	基于历史数据条件下，此航程的平均距离
<i>平均航程持续时间</i>	数字	基于历史数据条件下，此航程的平均持续时间
<i>平均航向</i>	数字	船舶的平均航向，以度为单位
<i>月份</i>	字符	当年的当前月份
<i>最佳完成距离</i>	数字	根据最佳路线完成航程的剩余距离
<i>最佳完成时间</i>	数字	根据最佳路线和历史航速完成航程的剩余时间
<i>船舶类型和尺寸</i>	字符	船舶尺寸和类型组别
<i>连线</i>	数字	此特定航程（即出发港到到达港）的标识

# 特征重要性

## ETA 模型的特征重要性

特征重要性是指特定特征在推导预测方面对模型的影响程度。

下表从影响最大到最小按降序列出了 ETA 模型的八个特征：

排名	特征名称	描述
1	<i>最佳完成时间</i>	根据最佳路线和历史航速完成航程的剩余时间
2	<i>最佳完成距离</i>	根据最佳路线完成航程的剩余距离
3	<i>平均航程距离</i>	基于历史数据条件下，此航程的平均距离
4	<i>船舶类型和尺寸</i>	船舶的类型，例如散装货船、集装箱船、运输船等
5	<i>连线</i>	此特定航程（即出发港到到达港）的标识
6	<i>到达港</i>	预测目的地模型中的预测目的地
7	<i>平均航向</i>	船舶的平均航向，以度为单位
8	<i>出发港</i>	开始航行的港口的名称

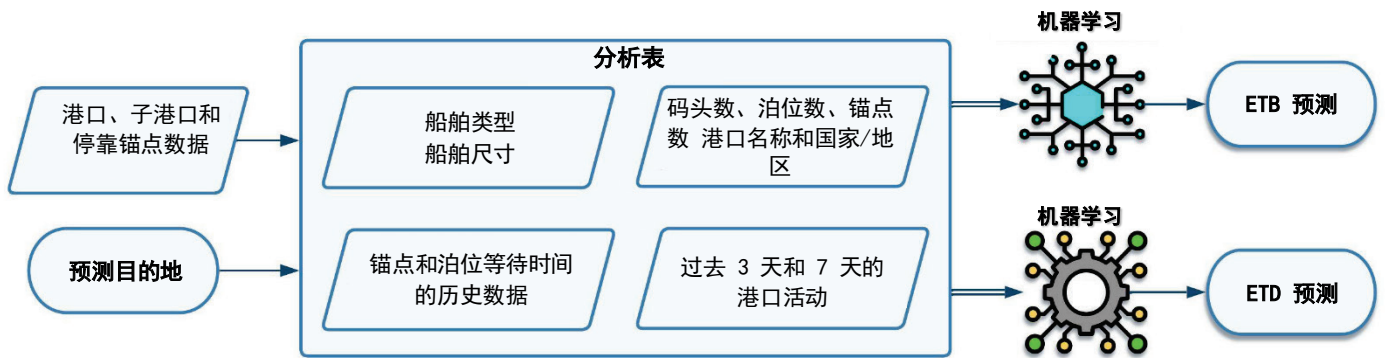
# 港口拥塞模型

# 概述

## 港口拥塞模型的简要概述

港口拥塞定义为抵达港口的船舶速率大于离开同一港口的船舶速率时的状况。

船队预测分析提供了对预测锚点等待时间、预计停泊时间 (ETB)、预计港口等待时间和预计离港时间 (ETD) 的估计。这种见解是在精细级别提供的 - 针对感兴趣的特定船舶类型和尺寸类别。

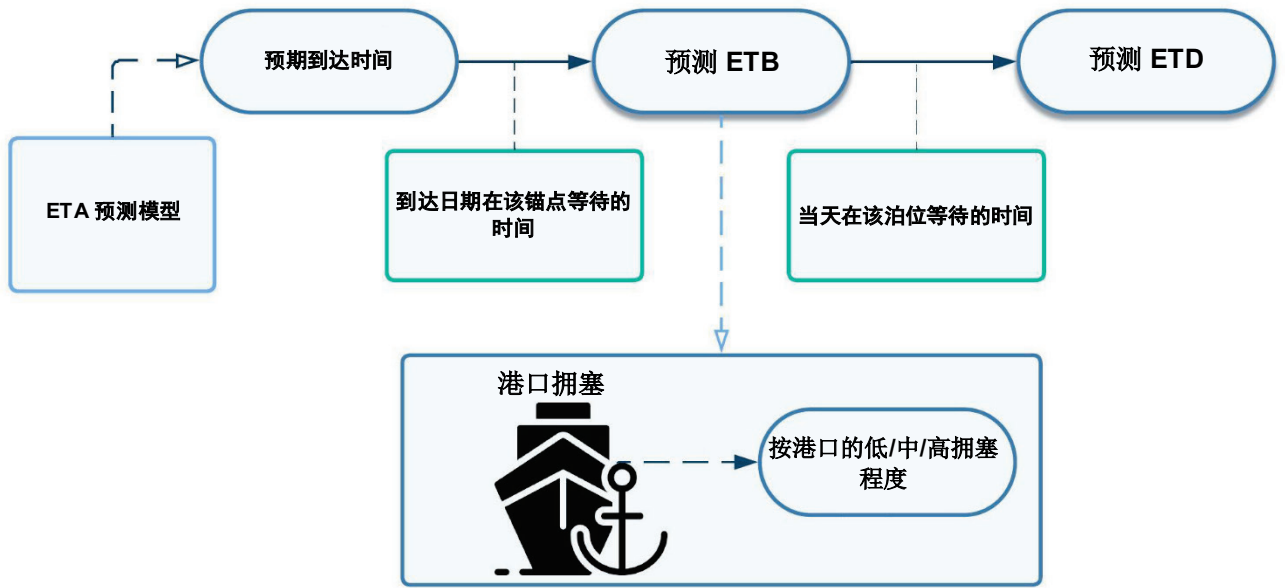


港口拥塞模型概述，其中包括 ETB 和 ETD 预测模型

## ETA、ETB 和 ETD 预测

ETA、ETB 和 ETD 预测模型协同工作，用以估算船舶预期到达泊位和预期离开的日期和时间。

对船舶预测目的地进行 ETA 预测后，ETB 和 ETD 模型会考虑船舶在未来 5 天内的预测动态。如果到达预测目的地的预期时间超过 ETB 预测的五天窗口，或 ETB 超过 ETD 预测的五天窗口，则分配给第五天的预测拥塞等级当前值将分别用于估计船舶的 ETB 和 ETD。



模型的实时协同效应

# 港口拥塞等级

## 如何计算港口拥塞等级

根据当前和过去的等待时间计算每个港口的拥塞等级。此等级还适用于预测船舶到达该地点的“预测 ETB”数据。该等级反映了港口的拥塞级别：

港口拥塞等级	锚点平均等待时间（小时）
低	0 - 2
中	2 - 12
高	12+

根据船舶在其锚点处的平均等待时间，每天为每个港口分配港口拥塞级别。过去 5 天的港口拥塞级别是根据实际历史数据得出的，而未来 5 天的拥塞级别是根据模型的 ETB 预测得出的。



# 港口拥塞预测

## 出发、到达和预测到达

我们的历史数据使我们能够了解每个港口和子港口海域处每个船舶类型和尺寸组合的到达和出发情况。在我们的 Seasearcher 界面，默认情况下会显示过去 5 天的历史信息，但可以将其扩展为包括过去一周、一月、一年或自定义范围。

此外，针对所选港口以及船舶类型和尺寸的组合，可提供未来五天的预测到达次数。通过计算所选船舶类型和尺寸的船舶（其预测目的地为所选港口，ETA 针对该特定日）的数量，推导出每天的预测到达数量。

## 平均等待时间预测

如港口拥塞模型的概述部分所述，锚点和泊位等待时间的五天预测是根据模型的输出项进行估计的。

这些等待时间是每个港口中船舶类型和尺寸的每个组合的平均值。此外，还会显示过去五天的锚点和泊位等待时间历史信息。如上所述，此历史信息范围可扩展至过去一周、一月、一年或自定义范围。

## 预测到达数

在 Seasearcher 的地点选项卡中选择了港口或子港口后，预测到达数选项卡下会显示预计到达所选港口或子港口的所有船舶的列表。该列表包括预测目的地为所选港口的船舶。

此选项卡是对现有选项卡的补充，即“预期到达数”、“到达数”、“出发数”、“照准数”和“地点附近”，每个选项卡的详细信息如下所述：

- *预期到达数* - 包括根据 AIS 信息预定到达港口或子港口的所有船舶
- *到达数* - 包括已到达且当前位于港口或子港口的所有船舶
- *出发数* - 包括曾在港口或子港口停靠并已出发的所有船舶
- *照准数* - 包括与港口或子港口互动的所有船舶的 AIS 照准数
- *地点附近* - 包括位于港口或子港口附近预选距离的所有船舶数

# 预计停泊时间 (ETB) 模型

## ETB 时间计算方法概述

预计停泊时间 (ETB) 模型可预测船舶到达其预测目的地锚点和兼容泊位之间的时间。

此数字对应于船舶在锚点处的等待时间，并用于计算预期船舶进入兼容泊位的时间。这是通过将船舶的 ETA 预测和 ETB 预测相加来计算的。

该模型预测了 5 天前船舶在港口的 ETB 或在锚点的等待时间。ETB 模型利用更多已验证的数据为预测提供更丰富的信息，例如：

- 已验证的历史港口、子港口和停靠锚点数据
- 结构港口信息
- 船舶特征数据
- （所有船舶类型和尺寸的）历史锚点等待时间

# 模型特征

## 在 ETB 模型中输入的特征

“预测 ETB” 的梯度提升模型的特征（或输入项）包括港口名称和港口特征（例如泊位数和锚点数）、历史交通模式数据和一些船舶特征信息（例如船舶的类型和尺寸）。

下表提供了模型特征的细化视图以及简短描述：

特征名称	取值	描述
港口名称	字符	港口的名称
港口类型	字符	指示它是港口还是子港口
港口国家/地区	字符	港口所在的国家/地区
港口泊位数	整数	泊位数
港口码头数	整数	码头数
港口锚点数	整数	锚点数
日期	日期	预测日期（年-月-日）
港口最大船舶数	数字	基于历史数据的港口最大船舶数
港口频率级别	低/中/高	指示基于历史数据的港口交通级别
过去 x 天的港口活动	是/否	指示港口在过去 x 天是否处于活动状态，其中 x 为 7、30、90、365 天
工作日	是/否	指示是否为工作日（星期一至星期五）
船舶类型	字符	船舶类型类别，例如散装货船、普通货船、油轮、液化气体运输船
船舶尺寸	字符	船舶尺寸类别，例如大灵便型、灵便型、巴拿马型、巨型油轮、苏伊士型
到达船舶数	整数	到达泊位的船舶数
过去 x 天在锚点的平均时间	整数	过去 x 天的锚点平均等待时间，其中 x 为 1、3、7 天
拥塞	是/否	指示是否有任何锚点等待时间
锚点拥塞率	数字	每个港口的锚点等待时间与平均历史等待时间的比率
港口 ID	数字	唯一的港口标识符
港口停靠排名	数字	基于历史港口停靠数的港口全球排名

# 特征重要性

## ETB 模型中的特征及其重要性加权

特征按照从影响最大到最小的顺序加权：

排名	特征名称	描述
1	过去 7 天在锚点的平均时间	过去几天的平均锚点等待时间
2	锚点拥塞率	每个港口的锚点等待时间与平均历史等待时间的比率
3	过去 3 天在锚点的平均时间	过去几天的平均锚点等待时间
4	船舶类型	船舶类型，例如散装货船、普通货船、油轮、液化天然气运输船
5	船舶尺寸	指示船舶尺寸，例如大灵便型、灵便型、巴拿马型、巨型油轮、苏伊士型
6	过去一天在锚点的平均时间	过去一天的平均锚点等待时间
7	港口国家/地区	港口所在的国家/地区
8	港口停靠排名	基于历史港口停靠数的港口全球排名
9	港口最大船舶数	此港口记录的历史最大船舶数
10	工作日	指示是否为工作日（星期一至星期五）

# 预计离港时间 (ETD) 模型

## 关于预测 ETD 模型工作原理的概述

预计离港时间 (ETD) 模型用于预测船舶将从港口出发的时间。

该模型可预测船舶的 ETD，并将其与预测的 ETA 和 ETB 一起提供。ETD 模型利用已验证的历史港口、子港口和停靠锚点数据，以及每个港口所有类型和尺寸的船舶到达泊位和停泊的历史等待时间。

# 模型特征

## ETD 模型中使用的特征

预测 ETD 预测中使用的特征（或输入项）包括港口名称和港口特征（例如泊位数和锚点数）、历史交通模式（2019 年 1 月以来）和一些船舶特征信息（例如船舶类型和尺寸）。

下表提供了模型特征的细化视图以及简短描述：

特征名称	取值	描述
港口名称	字符	港口的名称
港口类型	字符	指示它是港口还是子港口
港口国家/地区	字符	港口所在的国家/地区
港口泊位数	整数	泊位数
港口码头数	整数	码头数
港口锚点数	整数	锚点数
日期	日期	预测日期（年-月-日）
港口最大船舶数	数字	基于历史数据的港口最大船舶数
港口频率级别	低/中/高	指示基于历史数据的港口交通级别
过去 x 天的港口活动	是/否	指示港口在过去 x 天是否处于活动状态，其中 x 为 7、30、90、365 天
工作日	是/否	指示是否为工作日（星期一至星期五）
船舶类型	字符	船舶类型类别，例如散装货船、普通货船、油轮、液化气体运输船
船舶尺寸	字符	船舶尺寸类别，例如大灵便型、灵便型、巴拿马型、巨型油轮、苏伊士型
到达船舶数	整数	到达泊位的船舶数
过去 x 天在泊位的平均时间	整数	过去 x 天的停泊平均等待时间，其中 x 为 1、3、7 天
拥塞	是/否	指示是否有任何锚点等待时间
锚点拥塞率	数字	每个港口的锚点等待时间与平均历史等待时间的比率
港口 ID	数字	唯一的港口标识符
港口停靠排名	数字	基于历史港口停靠数的港口全球排名



# 特征重要性

## ETD 模型的特征重要性

下表从影响最大到最小按降序列出了 ETD 模型的九个特征：

排名	特征名称	描述
1	过去 7 天在泊位的平均时间	过去 7 天的平均停泊等待时间
2	过去一天在泊位的平均时间	过去一天的平均停泊等待时间
3	船舶尺寸	指示船舶尺寸，例如大灵便型、灵便型、巴拿马型、巨型油轮、苏伊士型
4	船舶类型	指示船舶类型，例如散装货船、普通货船、油轮、液化气体运输船
5	锚点拥塞	每个港口的锚点等待时间与平均历史等待时间的比率
6	港口国家/地区	港口所在的国家/地区
7	港口泊位数	此港口中的泊位数
8	过去 3 天在泊位的平均时间	过去 3 天的平均停泊等待时间
9	工作日	指示是否为工作日
10	港口停靠排名	基于历史港口停靠数的港口全球排名