**外贸业务需求爬虫分析报告**

风控部研究项目

本文对爬虫抓取的外贸论坛标题进行文本数据挖掘，分析了外贸企业的业务需求。

2014

卿聪

风控部

2014/9/25

目录

[项目目的 2](#_Toc399749789)

[研究方法 2](#_Toc399749790)

[出口企业需求分析 2](#_Toc399749791)

[1. 数据抓取 2](#_Toc399749792)

[2. 高频词统计 2](#_Toc399749793)

[3. 关联分析 4](#_Toc399749794)

[4. 分析和建议 7](#_Toc399749795)

[进口企业需求分析 8](#_Toc399749796)

[1. 数据抓取 8](#_Toc399749797)

[2. 高频词统计 8](#_Toc399749798)

[3. 关联分析 10](#_Toc399749799)

[4. 分析和建议 13](#_Toc399749800)

**外贸业务需求爬虫分析报告**

# 项目目的

通过爬虫抓取网络论坛上外贸企业间直接交流的文字信息进行分析，了解外贸企业实际需求，为我们的营销和服务提供方向和建议。

# 研究方法

本文采用定量和定性相结合的方法，分别对出口企业和进口企业进行分析。具体来说，本文利用网络爬虫从外贸论坛上抓取帖子标题，经筛选后对标题进行高频词统计，并对高频词进行关联分析，从而找到外贸企业的需求和问题。

# 出口企业需求分析

## 数据抓取

我们利用爬虫软件，从福步外贸论坛“出口交流”板块抓取前10000个标题。剔除重复的标题（重复发帖），剩下7266条。作为对照，我们选取其中包含问号的2311条标题，以排除广告贴等无关内容，确定找出来的高频词确实反映企业的问题和需求。

## 高频词统计

我们对这些标题进行高频词统计来确定外贸企业的需求。我们分别按2字词、3字词和4字词进行高频词统计，各保留频率最高的20个词。去掉如“怎么”、“问题”、“为什么”等无明显指向意义的虚词后，每个样本各得到有效高频词共计60个。统计结果如下表：

表1： 所有标题（7266条）：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **2字高频词** | **次数** | **3字高频词** | **次数** | **4字高频词** | **次数** |
| 客户 | 1828 | FOB | 149 | 付款方式 | 139 |
| 客人 | 383 | 信用证 | 131 | 客户要求 | 72 |
| 公司 | 351 | 目的港 | 47 | 外贸公司 | 64 |
| 付款 | 321 | CIF | 46 | 怎么操作 | 47 |
| 产品 | 240 | 土耳其 | 42 | 国外客户 | 44 |
| 样品 | 222 | L/C | 38 | 指定货代 | 42 |
| 货物 | 194 | 俄罗斯 | 35 | 尼日利亚 | 41 |
| 价格 | 181 | 寄样品 | 35 | 怎么回复 | 41 |
| 工厂 | 166 | 出口权 | 32 | 出口退税 | 38 |
| 提单 | 163 | 孟加拉 | 32 | 印度客户 | 32 |
| 邮件 | 153 | 供应商 | 31 | 该怎么回 | 29 |
| 退税 | 142 | 找客户 | 31 | 开发客户 | 29 |
| 货代 | 137 | 样品费 | 31 | FOB价 | 27 |
| 报关 | 131 | 船公司 | 28 | 巴基斯坦 | 27 |
| 清关 | 107 | 集装箱 | 28 | 出口货物 | 27 |
| 美国 | 104 | 增值税 | 28 | 客户怎么 | 27 |
| 海关 | 95 | 意大利 | 27 | 贸易公司 | 27 |
| 印度 | 94 | DHL | 25 | 怎么跟进 | 26 |
| 费用 | 92 | 是骗子 | 24 | 美国客户 | 23 |
| 订单 | 90 | 人民币 | 23 | 买单出口 | 22 |

表2： 含问号标题（2311条）：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **2字高频词** | **次数** | **3字高频词** | **次数** | **4字高频词** | **次数** |
| 客户 | 662 | FOB | 52 | 付款方式 | 47 |
| 公司 | 140 | 信用证 | 45 | 怎么操作 | 28 |
| 客人 | 132 | 目的港 | 18 | 客户要求 | 22 |
| 付款 | 97 | 船公司 | 16 | 外贸公司 | 22 |
| 样品 | 85 | 供应商 | 16 | 怎么回复 | 22 |
| 回复 | 76 | CIF | 15 | 国外客户 | 15 |
| 工厂 | 72 | L/C | 15 | 如何操作 | 15 |
| 邮件 | 71 | 集装箱 | 13 | 怎么跟进 | 15 |
| 产品 | 68 | 是骗子 | 13 | 指定货代 | 15 |
| 价格 | 65 | 样品费 | 13 | 巴基斯坦 | 12 |
| 报价 | 63 | 出口权 | 12 | 出口退税 | 12 |
| 操作 | 63 | 寄样品 | 12 | 贸易公司 | 11 |
| 货物 | 62 | 土耳其 | 12 | 尼日利亚 | 11 |
| 货代 | 59 | DHL | 11 | 怎么出口 | 11 |
| 提单 | 59 | 意大利 | 11 | 买单出口 | 10 |
| 退税 | 51 | 俄罗斯 | 10 | 开发客户 | 8 |
| 信用 | 49 | 增值税 | 10 | 美国客户 | 8 |
| 报关 | 42 | UPS | 9 | 货怎么办 | 7 |
| 费用 | 38 | 不一致 | 9 | 进出口权 | 7 |
| 开发 | 32 | 产地证 | 9 | 进口关税 | 7 |

不难发现，在两张表中，“客户”都是出现最多的词，频率远高于其他关键词。可以说，出口企业作为销售商是**以客户为导向**的，其需求也主要来自于吸引客户下单、和客户沟通、满足客户提的要求等环节。

从高频词统计中能够发现一些有趣的现象，如表1里国家名出现次数超过30次的包括**美国、印度、俄罗斯、尼日利亚、孟加拉**。其中，前三个是出口热门地，而通过对含尼日利亚和孟加拉的标题进行分析发现，尼日利亚之所以出现是因为该国的骗子多，而关于孟加拉的帖子很多是关于信用证问题的，因为该国付汇一般只能使用信用证。

由于全样本标题中包含外贸业务员交流经验、分享心得的帖子，所以表1更能全面反映外贸业务内容。而子样本都是包含问号的标题，所以表2更能反映外贸企业的问题和需求。而因为公司的业务和外贸企业的需求更相关，所以后面的分析我们主要采用含问号的子样本。

## 关联分析

为了寻找各关键词之间的关系以便做进一步分析，我们对关键词进行association分析，即寻找同一标题中出现某一关键词（或某几个关键词）时很可能同时出现的另一关键词。如发现标题中出现“回复”时，通常会同时出现“客户”和“邮件”，那么这些标题的意思很可能是“回复客户邮件”。这样，我们可以对关键词进行分类，获得了比原来的关键词更有用的信息。

首先，定义association分析中使用的术语如下表：

表3：

|  |  |  |
| --- | --- | --- |
| 名称 | 定义 | 公式 |
| X, Y | 关键词 |  |
| Support(X) | 所有标题中关键词X出现概率 | 含X的标题数/总标题数 |
| Confidence(X, Y) | 含X标题中出现Y的概率 | Support(X∩Y) / Support(X) |
| Lift(X, Y) | 标题含X后，Y出现概率提升倍数 | Confidence(X, Y) / Support(Y) |

注：Confidence(X, Y)和Lift(X, Y)越大，则关键词X和Y的关联性越强。若X和Y无关，则Lift(X, Y)=1。

我们对含问号的子样本进行关联分析，剔除掉明显包含关系的关键词组合（如出现“信用证”则必出现“信用”，confidence(信用证，信用)=1，反之则不然）。我们对confidence和lift指标设定阈值进行筛选（注：由于“客户”出现频率很高，所以confidence(X, 客户)普遍较大，而lift(X, 客户)较小，需鉴别），并按X分类整理得到关联规则如下表：

表4：

|  |  |  |  |
| --- | --- | --- | --- |
| X | Y | confidence | lift |
| {怎么回复} | {客户} | 0.41 | 0.96 |
| {邮件} | 0.18 | 3.92 |
| {怎么跟进} | {客户} | 0.93 | 2.19 |
| {邮件} | {客户} | 0.42 | 0.98 |
| {回复} | 0.22 | 4.31 |
| {样品} | {客户} | 0.57 | 1.33 |
| {信用} | {信用证} | 0.92 | 30.10 |
| {付款} | 0.10 | 1.64 |
| {外贸公司} | {工厂} | 0.41 | 8.57 |
| {退税} | {出口退税} | 0.24 | 28.90 |
| {报关} | 0.12 | 4.34 |
| {操作} | 0.10 | 2.29 |
| {公司} | 0.10 | 1.11 |
| {提单} | {货代} | 0.11 | 2.72 |
| {目的港} | {客户} | 0.29 | 0.69 |
| {客人} | 0.24 | 2.68 |
| {贸易公司} | {产品} | 0.36 | 8.08 |
| {客人} | {回复} | 0.10 | 1.97 |
| {客户要} | {客户要求} | 0.39 | 25.80 |
| {客户，邮件} | {回复} | 0.21 | 4.13 |
| {开发} | {客户} | 0.47 | 1.10 |
| {开发客户} | 0.25 | 45.16 |
| {价格} | {客户} | 0.40 | 0.94 |
| {FOB} | 0.12 | 3.31 |
| {货代} | {指定货代} | 0.27 | 25.80 |
| {FOB} | 0.13 | 3.54 |
| {工厂} | 0.13 | 2.62 |
| {费用} | 0.11 | 4.07 |
| {提单} | 0.11 | 2.72 |
| {货代，客户} | {指定货代} | 0.45 | 43.79 |
| {回复} | {怎么回复} | 0.29 | 19.27 |
| {邮件} | 0.20 | 4.31 |
| {客人} | 0.17 | 1.97 |
| {报价} | 0.11 | 2.61 |
| {供应商} | {客户} | 0.56 | 1.32 |
| {公司} | {外贸公司} | 0.17 | 11.12 |
| {船公司} | 0.11 | 11.12 |
| {工厂} | 0.11 | 2.26 |
| {公司，客户} | {外贸公司} | 0.17 | 11.32 |
| {工厂} | {公司} | 0.20 | 2.26 |
| {外贸公司} | 0.13 | 8.57 |
| {货代} | 0.10 | 2.62 |
| {工厂，公司} | {外贸公司} | 0.64 | 42.22 |
| {付款} | {付款方式} | 0.50 | 15.71 |
| {费用} | {FOB} | 0.16 | 4.47 |
| {货代} | 0.16 | 4.07 |
| {出口到} | {货物} | 0.11 | 2.52 |
| {产品，公司} | {贸易公司} | 0.57 | 75.06 |
| {操作} | {怎么操作} | 0.44 | 22.94 |
| {如何操作} | 0.24 | 22.94 |
| {报价} | {客户} | 0.27 | 0.64 |
| {客人} | 0.15 | 1.74 |
| {回复} | 0.14 | 2.61 |
| {报关} | {退税} | 0.15 | 4.34 |
| {货物} | 0.10 | 2.33 |
| {付款} | 0.10 | 1.57 |
| {FOB} | {价格} | 0.14 | 3.31 |
| {货代} | 0.14 | 3.54 |
| {费用} | 0.12 | 4.47 |

注：包含关系的词组如“信用、信用证”的lift值很高，不具参考意义。通常lift在2到10之间的关联关系较有参考价值。

将高频词按照association分析结果分类(X, Y成对出现的分成一类)，字体染成同一个颜色，得到下表：

表5：

|  |  |  |
| --- | --- | --- |
| 2字高频词 | 3字高频词 | 4字高频词 |
| 客户 | FOB | **付款方式** |
| 公司 | **信用证** | **怎么操作** |
| 客人 | 目的港 | 客户要求 |
| 付款 | **船公司** | 外贸公司 |
| 样品 | 供应商 | 怎么回复 |
| 回复 | CIF | 国外客户 |
| 工厂 | **L/C** | 如何操作 |
| 邮件 | 集装箱 | 怎么跟进 |
| 产品 | 是骗子 | 指定货代 |
| 价格 | 样品费 | 巴基斯坦 |
| 报价 | 出口权 | **出口退税** |
| 操作 | 寄样品 | 贸易公司 |
| 货物 | 土耳其 | 尼日利亚 |
| 货代 | DHL | 怎么出口 |
| 提单 | 意大利 | 买单出口 |
| 退税 | 俄罗斯 | **开发客户** |
| 信用 | **增值税** | 美国客户 |
| 报关 | UPS | 货怎么办 |
| 费用 | 不一致 | 进出口权 |
| 开发 | 产地证 | 进口关税 |

注：同一个词可能属于几个类中，如“报价”既可以是针对产品，又可以是和客户沟通过程中发邮件报价。

根据表格，直接按association分析的结果，排在前面的高频词可以分为以下几类：

一、客户（客人）、开发客户、客户要求等：这一类是出口外贸企业业务需求的主体，即如何开发客户、满足客户需求。

二、回复、邮件、报价等：这一类是出口企业和客户沟通过程中的问题，可以看出邮件是交流的主要工具。

三、货物、报关、（出口）退税、操作等：这一类是订单下好后出口过程中报关等流程中遇到的问题。

四、货代、提单、费用、FOB等：这一类是报关后运输过程中遇到的问题。

五、付款（方式）、信用证、L/C等：这一类是出口企业收款过程中遇到的问题。

2字高频词中未被association分类的如下：公司、工厂、产品、价格、样品。可以看出，产品、价格、样品、报价可以分作一类，表示出口企业自身产品存在的一些问题。而公司、外贸公司、工厂在很多情况下是外贸企业自称，看不出明显需求，故不分类。

综上所述，可以把出口企业需求分类成下表：

表6：

|  |  |  |
| --- | --- | --- |
| **需求分类** | **关键词** | **关键词频数比例** |
| **客户** | 客户、客人、开发客户、客户要求 | 42.1% |
| **产品** | 产品、价格、样品、报价 | 14.9% |
| **运输** | 货代、提单、费用、FOB、船公司 | 11.9% |
| **报关** | 货物、报关、退税、操作 | 11.6% |
| **和客户沟通** | 回复、邮件、报价 | 11.1% |
| **收款** | 付款（方式）、信用证、L/C | 8.3% |

## 4. 分析和建议

一、出口企业以客户为导向，其难点和问题主要来自于开发客户、与客户沟通、满足客户需求。我们在为出口企业服务的同时应考虑到客户的客户（即海外买家）的信息，帮助买卖双方解决好出口流程中产生的问题。

二、操作性问题包括报关和运输。我们应以专业操作为客户节省时间和费用。特别是退税方面，我们应利用外贸综合服务平台优势，开发出口退税业务。

三、收款环节，出口企业关心的是减小收汇风险。我们应帮助出口企业合理选择付款方式，特别是信用证方面帮助他们合理审证，减少风险。

四、产品方面，我们应增强对各类产品熟悉度，成为产品行家。同时，可以利用平台进行产品信息发布等活动。

# 进口企业需求分析

## 数据抓取

从福步外贸论坛“进口交流”板块抓取标题5205个。剔除重复标题后还剩下5145条。由于含问号的标题数量仅有241个，我们扩大范围至含“吗、？、问题、求助、求教、请教”等词的标题，去除重复项后得到1567个作为对照组。

## 高频词统计

按照和出口论坛同样的方法，得到下表：

表7： 所有标题（5145条）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **2字高频词** | **次数** | **3字高频词** | **次数** | **4字高频词** | **次数** |
| 公司 | 218 | 增值税 | 62 | 进口关税 | 80 |
| 货物 | 181 | 信用证 | 61 | 进口货物 | 79 |
| 关税 | 175 | 许可证 | 50 | 进料加工 | 56 |
| 操作 | 150 | 废塑料 | 48 | 怎么操作 | 51 |
| 产品 | 150 | 供应商 | 47 | 进口清关 | 50 |
| 清关 | 150 | 化妆品 | 44 | 进口付汇 | 47 |
| 加工 | 147 | 从美国 | 32 | 进口食品 | 47 |
| 报关 | 133 | 葡萄酒 | 25 | 来料加工 | 38 |
| 流程 | 129 | 铁矿石 | 25 | 什么手续 | 37 |
| 代理 | 128 | 报关单 | 24 | 进口代理 | 35 |
| 手续 | 118 | 产地证 | 24 | 进口许可 | 33 |
| 设备 | 112 | 进口税 | 24 | 进口报关 | 32 |
| 付汇 | 102 | 意大利 | 24 | 进口红酒 | 32 |
| 香港 | 102 | FOB | 20 | 请教进口 | 32 |
| 海关 | 100 | 进口奶 | 20 | 台湾进口 | 32 |
| 食品 | 100 | 橄榄油 | 19 | 美国进口 | 30 |
| 费用 | 98 | 从香港 | 18 | 代理进口 | 27 |
| 办理 | 97 | 到香港 | 18 | 进口流程 | 27 |
| 客户 | 96 | 电解铜 | 18 | 进口设备 | 27 |
| 美国 | 90 | 原产地 | 18 | 从美国进 | 26 |

表8： 子样本（1567条）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **2字高频词** | **次数** | **3字高频词** | **次数** | **4字高频词** | **次数** |
| 关税 | 80 | 增值税 | 25 | 进口关税 | 37 |
| 加工 | 60 | 信用证 | 22 | 进口货物 | 30 |
| 货物 | 56 | 许可证 | 21 | 进料加工 | 29 |
| 公司 | 49 | 进口商 | 12 | 进口付汇 | 23 |
| 付汇 | 45 | 报关单 | 11 | 进口清关 | 20 |
| 设备 | 45 | 化妆品 | 11 | 来料加工 | 20 |
| 清关 | 42 | 从美国 | 10 | 进口食品 | 19 |
| 报关 | 41 | 香港进 | 10 | 关税问题 | 18 |
| 操作 | 36 | FOB | 9 | 进口核销 | 15 |
| 进料 | 35 | 美国进 | 9 | 进口需要 | 15 |
| 付款 | 34 | CIF | 8 | 税的问题 | 15 |
| 海关 | 34 | 产地证 | 8 | 进口设备 | 14 |
| 核销 | 34 | 铁矿石 | 8 | 清关问题 | 12 |
| 产品 | 33 | 从香港 | 7 | 付款问题 | 11 |
| 食品 | 33 | 到香港 | 7 | 核销问题 | 11 |
| 香港 | 33 | 废塑料 | 7 | 进口许可 | 11 |
| 贸易 | 32 | 供应商 | 7 | 单的问题 | 10 |
| 手续 | 32 | 进口税 | 7 | 加工贸易 | 10 |
| 代理 | 27 | 保健品 | 6 | 进口报关 | 9 |
| 费用 | 26 | 从日本 | 6 | 进口代理 | 9 |

容易发现，在出口论坛标题中占主导地位的关键词“客户”在进口论坛中出现次数很少，而一些产品名却经常出现。如果说出口企业需求来源是以人（客户）为中心的话，那么进口企业需求来源则是**以物为中心**。

从高频词统计中，同样可以发现一些有趣的规律。如出现频率较高的国家和地区依次是**香港、美国、日本、台湾、德国、意大利、印度、印尼**。和出口相比，进口热门国更多的是发达国家，这些国家的产品品质更好。而热门产品包括**设备**（机电、机器、机械、汽车）、**食品**（葡萄酒、奶粉、红酒、橄榄油、保健品）、**原料**（废塑料、木材、铁矿石、电解铜）、**化妆品**等。

## 关联分析

同样，我们对子样本高频词进行association分析，整理得到下表：

表9：

|  |  |  |  |
| --- | --- | --- | --- |
| X | Y | confidence | lift |
| {报关,付汇} | {核销} | 0.50 | 13.78 |
| {报关,核销} | {付汇} | 0.63 | 13.64 |
| {报关} | {进口报关} | 0.25 | 16.95 |
| {代理,公司} | {进口代理} | 0.26 | 16.43 |
| {代理} | {进口代理} | 0.28 | 17.35 |
| {公司} | 0.18 | 1.92 |
| {费用} | {清关} | 0.09 | 1.43 |
| {手续} | 0.08 | 1.57 |
| {付汇,核销} | {报关} | 0.21 | 3.53 |
| {付汇} | {进口付汇} | 0.47 | 21.82 |
| {核销} | 0.24 | 6.55 |
| {报关} | 0.10 | 1.68 |
| {操作} | 0.09 | 1.33 |
| {付款} | {付款问题} | 0.14 | 29.00 |
| {公司,货物} | {进口货物} | 0.42 | 11.77 |
| {公司} | {贸易} | 0.11 | 1.78 |
| {代理} | 0.11 | 1.92 |
| {关税} | {进口关税} | 0.45 | 12.67 |
| {增值税} | 0.13 | 4.86 |
| {关税问题} | 0.11 | 12.67 |
| {关税问题} | {进口关税} | 0.47 | 13.22 |
| {关问题} | {清关} | 0.39 | 5.84 |
| {报关} | 0.16 | 2.73 |
| {海关} | {货物} | 0.10 | 1.24 |
| {核销} | {进口核销} | 0.31 | 27.55 |
| {付汇} | 0.30 | 6.55 |
| {进口付汇} | 0.18 | 8.21 |
| {核销问题} | 0.14 | 27.55 |
| {报关} | 0.10 | 1.70 |
| {货物} | {进口货物} | 0.44 | 12.38 |
| {加工} | {进料加工} | 0.43 | 17.09 |
| {来料加工} | 0.29 | 17.09 |
| {贸易} | 0.21 | 3.37 |
| {加工贸易} | 0.20 | 17.09 |
| {进口代理} | {公司} | 0.17 | 1.82 |
| {进口付汇} | {核销} | 0.30 | 8.21 |
| {进口关税} | {增值税} | 0.14 | 5.11 |
| {进口关税} | {关税问题} | 0.11 | 13.22 |
| {进口核销} | {核销问题} | 0.28 | 56.10 |
| {进口清关} | {清关问题} | 0.16 | 29.39 |
| {进口许可} | {许可证} | 0.84 | 37.95 |
| {进料} | {进料加工} | 0.93 | 36.73 |
| {来料加工} | 0.17 | 9.67 |
| {贸易} | 0.08 | 1.34 |
| {进料加工} | {来料加工} | 0.18 | 10.36 |
| {来料加工} | {进料加工} | 0.26 | 10.36 |
| {贸易} | 0.13 | 2.12 |
| {贸易} | {加工} | 0.20 | 3.37 |
| {加工贸易} | 0.19 | 16.09 |
| {公司} | 0.17 | 1.78 |
| {清关} | {进口清关} | 0.34 | 15.10 |
| {清关问题} | 0.08 | 15.10 |
| {设备} | {进口设备} | 0.24 | 19.68 |
| {食品} | {进口食品} | 0.47 | 22.26 |
| {税问题} | {关税} | 0.49 | 6.18 |
| {增值税} | 0.12 | 4.48 |
| {香港} | {货物} | 0.14 | 1.69 |
| {公司} | 0.11 | 1.12 |
| {许可证} | {进口许可} | 0.55 | 37.95 |
| {增值税} | {关税} | 0.38 | 4.86 |
| {进口关税} | 0.18 | 5.11 |

将高频词按照association分析结果分类得到下表：

表10：

|  |  |  |
| --- | --- | --- |
| 2字高频词 | 3字高频词 | 4字高频词 |
| 关税 | **增值税** | **进口关税** |
| 加工 | 信用证 | 进口货物 |
| 货物 | 许可证 | **进料加工** |
| 公司 | 进口商 | **进口付汇** |
| 付汇 | **报关单** | 进口清关 |
| 设备 | 化妆品 | **来料加工** |
| 清关 | 从美国 | 进口食品 |
| 报关 | **香港进** | **关税问题** |
| 操作 | FOB | **进口核销** |
| 进料 | 美国进 | 进口需要 |
| 付款 | CIF | **税的问题** |
| 海关 | 产地证 | 进口设备 |
| 核销 | 铁矿石 | 清关问题 |
| 产品 | **从香港** | 付款问题 |
| 食品 | **到香港** | **核销问题** |
| 香港 | 废塑料 | 进口许可 |
| 贸易 | 供应商 | 单的问题 |
| 手续 | **进口税** | 加工贸易 |
| 代理 | 保健品 | 进口报关 |
| 费用 | 从日本 | 进口代理 |

从association分析结果来看，高频词中明显有几类可以分出来：

一、关税、增值税等：这一类是税务问题，包括查询具体产品的税率和避税等。

二、加工、进料、贸易、来料加工、进料加工等：这一类是加工贸易问题。

三、核销、付汇、报关等：这一类是交易完成后进口企业向海关申请报核的流程。

四、香港、公司、货物等：这一类是很多企业进口时先把货物运到香港再转运内地来获得关税优惠，或在香港设立离岸公司方便进口。

五、清关、费用、手续：这一类是进口企业清关过程中遇到的问题。

剩下未被分类的高频词中，我们认为设备、产品、食品等属于具体产品类，付款和信用证等属于付款类，代理、公司、进口代理、许可证、进口许可等属于进口代理类。

综上所述，我们把进口企业需求分类整理成下表：

表11：

|  |  |  |
| --- | --- | --- |
| **需求分类** | **关键词** | **关键词频数比例** |
| **产品** | 产品、食品、设备 | 15.4% |
| **税** | 关税、增值税 | 14.6% |
| **清关** | 清关、费用、手续、海关、操作 | 14.4% |
| **加工贸易问题** | 加工、贸易、来料加工、进料加工 | 14.0% |
| **核销** | 核销、付汇、报关 | 13.8% |
| **代理** | 代理、公司、许可证、进口许可 | 10.5% |
| **香港相关问题** | 香港、公司、货物 | 9.5% |
| **付款** | 付款、信用证 | 7.8% |

不难发现，和出口相比，进口企业需求的分类更多，各类的比例也更接近，这也说明进口流程的环节更多更杂，可操作的空间也更大。

## 4. 分析和建议

一、关税和增值税是进口企业关注的重点问题。我们应加强对税务技能和知识的掌握，为客户提供专业解答和税金垫付等服务。

二、操作方面包括以代理为基础的报关、清关、核销等流程。我们应该利用一站式外贸服务平台的专业优势，为客户节省时间和费用。

三、加工贸易是进口业务的重要类别。和一般贸易相比，其在税务、资质、核销等方面都有较大区别。我们应熟悉加工贸易实务，了解其风险点。

四、付款方面，我们应该帮助客户选择安全、节省的付款方式，特别是加强对信用证资源的运用，发展代开证业务。

五、产品方面，我们重点关注食品、设备等热门产品，了解相关风险和政策，拓展业务方向。

六、香港是中国唯一的自由贸易港，在税率、手续、外汇管制方面比内地有优势。很多进口企业选择香港转口贸易或设立香港离岸公司进行操作。由于经香港进口的货物常涉及包税进口等问题，我们应加强对于此类风险的防范。